Online Appendix: Regulation Design in Insurance Markets

Dhruva Bhaskar[*], Andrew McClellan[†], and Evan Sadler[‡]

June 8, 2023

1 Additional Proofs

Proofs of Propositions 3 and 4

We first prove Proposition 4. Proposition 3 is a corollary of Proposition 4 and its proof. To see this, take the existing allocation to be the laissez-faire outcome and the existing regulatory policy is exactly the set menus the firm offers after each signal if unrestricted—there are no off-path contracts in the existing policy. In reality, a laissez-faire policy would allow the firm to offer any menu of contracts, but for the purposes of our exercise, one should imagine a regulatory policy that restricts the firm to exactly those menus it is offering on the market. This makes finding local improvements maximally difficult. Proposition 4 states that we can find a local improvement, and the proof below constructs the improvement using two added latent contracts for each signal.

In what follows, we describe contracts as a pair (z_0, z_1) in which z_i are net transfers to the agent in state ω_i—namely, $z_i = -p + t(\omega_i)$. To avoid clutter, we drop dependence on x in U and u.

Let $\mathcal{R}$ be a regulatory policy implementing $\mathbf{a}$, with $M^s \in \mathcal{R}$ the menu the firm offers after signal s. Without loss, assume $\mathcal{R} = \{M^s\}_{s\in S}$.[1] We further assume the menu $M^s = \{(z_0^1, z_1^1), ..., (z_0^{N_s}, z_1^{N_s})\}$ is ordered so that $z_0^n > z_0^{n+1}$ and $z_1^n < z_1^{n+1}$—the net transfer in ω_0 is decreasing, and the net transfer in ω_1 is increasing. This is without loss as any contract that offers a lower net transfer in both events than another is never chosen by any type of agent either on or off path. Similarly, we assume without loss that every contract in M^s would be

[*]Baruch College, CUNY
[†]University of Chicago, Booth School of Business
[‡]Columbia University
[1]Adding menus to the policy only makes the exercise easier.

chosen by some type θ—not necessarily in the support of s—if offered the menu M^s. If not, we can remove those contracts, and the policy we construct still implements $\mathbf{a}$.

Within the implementation of $\mathbf{a}$ via $\mathcal{R}$, for each M^s, let c^s_θ denote the contract type θ chooses if offered M^s. For any contract $c \in M^s$, denote the set of types not in the support of s that choose c from M^s as

$$\Theta^s(c) := \{\theta \in \Theta \backslash supp(s) | c^s_\theta = c\}.$$

To reduce notational clutter, we will write $\Theta^s((z_0, z_1))$ as $\Theta^s(z_0, z_1)$. We introduce further useful notation:

- Let $\overline{u}_z = \max\{u_z(\omega_i, z_i)|\ i \in \{0, 1\},\ (z_0, z_1) \in \mathcal{C}\}$ be the highest marginal utility in any event of any contract in $\mathcal{C}$.
- Let $\underline{\mu} = \min\{\mu(\theta|s)|\ s \in S, \theta \in supp(s)\}$.
- Let $\underline{\theta} = \min\{\theta \in \Theta\}$ and $\overline{\theta} = \max\{\theta \in \Theta\}$.
- Let $\underline{\Delta}_\Theta = \min\{|\theta - \theta'| :\ \ \theta, \theta' \in \Theta,\ \theta' \neq \theta\}$ and $\overline{\Delta}_\Theta = \overline{\theta} - \underline{\theta}$.

Outline of the proof: To obtain a new regulatory policy that implements an improvement, we take each menu M^s, modify the existing contracts, and add up to two additional contracts. Both the modified contracts and the new contracts are close to contracts already in use. As we have a target allocation $\hat{\mathbf{a}}$ in mind, we necessarily replace each contract in M^s intended for some type $\theta \in supp(s)$ with the corresponding contract in the allocation $\hat{\mathbf{a}}$. Each contract in M^s that is chosen only by types $\theta \notin supp(s)$ is replaced with a modified version that still attracts type θ given the other modified contracts. Nevertheless, moving to $\hat{\mathbf{a}}$ from $\mathbf{a}$ may require more deterrence power than these modified off-path contracts provide, so we construct two additional latent contracts and add them to the menu. These new contracts attract types not in the support of s that previously would choose the contracts with the highest and lowest respective transfers in ω_1.

We first prove a series of lemmas which we use to modify and construct contracts as described above while staying within required bounds. The first lemma shows that given any contract $c \in \mathcal{C}$ and some type θ, we can modify the contract continuously such that all types above θ prefer the modified contract to c, all types below θ prefer c to the modified contract, and the firm's expected profit from selling the modified contract to types above θ is worse than selling them c. Moreover, we provide bounds on the utility differences between the two contracts for any type.

Lemma 1. *There exists an unbounded and strictly increasing differentiable function $f(\alpha) \geq 0$, a continuous function $\Delta(z, \alpha, \theta)$, and a constant $k \geq 0$ such that, for any contract $z =$*

$(z_0, z_1) \in \mathcal{C}$ *that does not underinsure, any* θ, *and any* $\alpha \geq 0$ *we have*

$$
\begin{aligned}
&U(\theta', (z_0 - \Delta(z, \alpha, \theta), z_1 + \alpha)) - U(\theta', (z_0, z_1)) \in (f(\alpha), k\alpha) \ \forall \theta' \geq \theta, \\
&U(\theta'', (z_0 - \Delta(z, \alpha, \theta), z_1 + \alpha)) - U(\theta'', (z_0, z_1)) < -f(\alpha) \ \forall \theta'' < \theta, \\
&\Pi(\theta', (z_0 - \Delta(z, \alpha, \theta), z_1 + \alpha)) - \Pi(\theta', (z_0, z_1)) < -\frac{\underline{\Delta}_\Theta}{2}\alpha \ \forall \theta' \geq \theta.
\end{aligned}
$$

Moreover, $\Delta(z, \alpha, \theta) \in [0, \frac{\theta}{1-\theta}\alpha]$.

Proof. Let $\overline{z}_1 = \max_{(z_0', z_1') \in \mathcal{C}} z_1'$. Define $f(\alpha) = \frac{\underline{\Delta}_\Theta}{2}(u(\omega_1, \overline{z}_1 + \alpha) - u(\omega_1, \overline{z}_1))$ and $k = \frac{\overline{\Delta}_\Theta}{2(1-\overline{\theta})}\overline{u}_z$. That u is differentiable, increasing and unbounded in z implies f is an unbounded and strictly increasing differentiable function. Fix any θ and let $\theta_- = \max\{\theta' \in \Theta : \ \theta' < \theta\}$ and $\theta_c = \frac{\theta + \theta_-}{2}$. For each $\alpha \geq 0$, define $\Delta(z, \alpha, \theta)$ to satisfy

$$
u(\omega_0, z_0 - \Delta(z, \alpha, \theta)) = u(\omega_0, z_0) + \frac{\theta_c}{1 - \theta_c}[u(\omega_1, z_1) - u(\omega_1, z_1 + \alpha)]. \tag{1}
$$

Note that $\Delta(z, 0, \theta) = 0$. Taking the derivative with respect to α, we get

$$
\frac{\partial \Delta(z, \alpha, \theta)}{\partial \alpha} = \frac{\theta_c u_z(\omega_1, z_1 + \alpha)}{(1 - \theta_c) u_z(\omega_0, z_0 - \Delta(z, \alpha, \theta))} \leq \frac{\theta_c}{(1 - \theta_c)},
$$

where the inequality follows from the fact that, because (z_0, z_1) does not underinsure, we have $u_z(\omega_0, z_0) \geq u_z(\omega_1, z_1)$, and therefore, $u_z(\omega_0, z_0 - \Delta(z, \alpha, \theta)) \geq u_z(\omega_1, z_1 + \alpha)$. Thus, we have

$$
\Delta(z, \alpha, \theta) = \int_0^\alpha \frac{\partial \Delta(z, \alpha', \theta)}{\partial \alpha'} d\alpha' \leq \frac{\theta_c}{1 - \theta_c}\alpha \leq \frac{\theta}{1 - \theta}\alpha.
$$

For any θ', using (1) we have

$$
\begin{aligned}
&U(\theta', (z_0 - \Delta(z, \alpha, \theta), z_1 + \alpha)) - U(\theta', (z_0, z_1)) \qquad (2) \\
&= (1 - \theta')u(\omega_0, z_0 - \Delta(z, \alpha, \theta)) + \theta' u(\omega_1, z_1 + \alpha) - (1 - \theta')u(\omega_0, z_0) - \theta' u(\omega_1, z_1) \\
&= [u(\omega_1, z_1 + \alpha) - u(\omega_1, z_1)]\big(\theta' - (1 - \theta')\frac{\theta_c}{1 - \theta_c}\big) \\
&= \frac{\theta' - \theta_c}{1 - \theta_c}[u(\omega_1, z_1 + \alpha) - u(\omega_1, z_1)].
\end{aligned}
$$

If $\theta' < \theta_c$, then $\frac{\theta' - \theta_c}{1 - \theta_c} < \theta' - \theta_c \leq \frac{-\underline{\Delta}_\Theta}{2}$ and so (2) implies

$$
\begin{aligned}
U(\theta', (z_0 - \Delta(z, \alpha, \theta), z_1 + \alpha)) - U(\theta', (z_0, z_1)) &< \frac{-\underline{\Delta}_\Theta}{2}[u(\omega_1, z_1 + \alpha) - u(\omega_1, z_1)] \\
&\leq -f(\alpha),
\end{aligned}
$$

where the last inequality follows from the fact that, because u is concave in z, $u(\omega_1, z_1 + \alpha) - u(\omega_1, z_1) \geq u(\omega_1, \overline{z}_1 + \alpha) - u(\omega_1, \overline{z}_1)$. If $\theta' > \theta_c$, then $\frac{\theta' - \theta_c}{1-\theta_c} > \theta' - \theta_c \geq \frac{\underline{\Delta}_\Theta}{2}$ and so

$$\begin{aligned} U(\theta', (z_0 - \Delta(z, \alpha, \theta), z_1 + \alpha)) - U(\theta', (z_0, z_1)) &> \frac{\underline{\Delta}_\Theta}{2}[u(\omega_1, z_1 + \alpha) - u(\omega_1, z_1)] \\ &\geq f(\alpha), \end{aligned}$$

and also, because $\frac{\theta' - \theta_c}{1-\theta_c} < \frac{\overline{\Delta}_\Theta}{2(1-\overline{\theta})}$

$$\begin{aligned} U(\theta', (z_0 - \Delta(z, \alpha, \theta), z_1 + \alpha)) - U(\theta', (z_0, z_1)) &< \frac{\overline{\Delta}_\Theta}{2(1-\overline{\theta})}[u(\omega_1, z_1 + \alpha) - u(\omega_1, z_1)] \\ &= \frac{\overline{\Delta}_\Theta}{2(1-\overline{\theta})} \int_0^\alpha u_z(\omega_1, z_1 + a) da \\ &< \frac{\overline{\Delta}_\Theta}{2(1-\overline{\theta})} \overline{u}_z \alpha \\ &= k\alpha. \end{aligned}$$

Firm profits for $\theta' \geq \theta_c$ satisfy

$$\begin{aligned} \Pi(\theta', (z_0 - \Delta(z, \alpha, \theta), z_1 + \alpha)) - \Pi(\theta', (z_0, z_1)) &= \Delta(z, \alpha, \theta)(1 - \theta') - \alpha\theta' \\ &\leq \alpha((1 - \theta')\frac{\theta_c}{1-\theta_c} - \theta') \\ &\leq -\frac{\underline{\Delta}_\Theta}{2}\alpha. \end{aligned}$$

■

An analogous argument works in reverse for contracts that do not overinsure, only now we increase transfers in state ω_0 and decrease transfers in ω_1. We omit the proof as it is analogous to that for the previous lemma.

Lemma 2. *There exists an unbounded and strictly increasing differentiable function $g(\alpha) \geq 0$, a continuous function $\eta(z, \alpha, \theta)$, and a constant $k' \geq 0$ such that, for any contract $z = (z_0, z_1) \in \mathcal{C}$ that does not overinsure, any θ, and any $\alpha \geq 0$ we have*

$$\begin{aligned} &U(\theta', (z_0 + \alpha, z_1 - \eta(z, \alpha, \theta))) - U(\theta', (z_0, z_1)) \in (g(\alpha), k'\alpha) \ \forall \theta' \leq \theta \\ &U(\theta'', (z_0 + \alpha, z_1 - \eta(z, \alpha, \theta))) - U(\theta'', (z_0, z_1)) < -g(\alpha) \ \forall \theta'' > \theta, \\ &\Pi(\theta', (z_0 + \alpha, z_1 - \eta(z, \alpha, \theta))) - \Pi(\theta', (z_0, z_1)) < -\frac{\underline{\Delta}_\Theta}{2}\alpha \ \forall \theta' \leq \theta. \end{aligned}$$

Moreover, $\eta(z, \alpha, \theta) \in [0, \frac{1-\theta}{\theta}\alpha]$.

The following lemma bounds the utility difference between any two contracts for a fixed type as a function of the distance between the two contracts. Recall that $\overline{u}_z$ is the highest marginal utility achieved in any event by any contract in $\mathcal{C}$.

Lemma 3. *For all contracts* (z_0, z_1) *and* (z_0', z_1') *in* $\mathcal{C}$ *with* $||(z_0, z_1) - (z_0', z_1')|| \leq \epsilon$*, we have* $|U(\theta, (z_0, z_1)) - U(\theta, (z_0', z_1'))| \leq \overline{u}_z \epsilon$ *.*

Proof. Take (z_0, z_1) and (z_0', z_1') in $\mathcal{C}$ with $||(z_0, z_1) - (z_0', z_1')|| \leq \epsilon$ and note

$$\begin{aligned}
&|U(\theta, (z_0, z_1)) - U(\theta, (z_0', z_1'))| \\
&\qquad \leq U(\theta, \max\{z_0, z_0'\}, \max\{z_1, z_1'\}) - U(\theta, \min\{z_0, z_0'\}, \min\{z_1, z_1'\}) \\
&\qquad \leq U(\theta, \min\{z_0, z_0'\} + \epsilon, \min\{z_1, z_1'\} + \epsilon) - U(\theta, \min\{z_0, z_0'\}, \min\{z_1, z_1'\}) \\
&\qquad \leq (1-\theta) \int_0^\epsilon u_z(\omega_0, \min\{z_0, z_0'\} + y) dy + \theta \int_0^\epsilon u_z(\omega_1, \min\{z_1, z_1'\} + y) dy \\
&\qquad \leq \epsilon \overline{u}_z.
\end{aligned}$$

■

The next lemma establishes a single-crossing property. If contract c has a higher net transfer than contract c' in ω_1, and a lower net transfer in ω_0, and some type θ prefers c to c', then all higher types also prefer c to c'. Analogously, if some type θ prefers c' to c, then all lower types also prefer c' to c.

Lemma 4. *Suppose* (z_0, z_1) *and* (z_0', z_1') *are contracts with* $z_1 < z_1'$ *and* $z_0' < z_0$*. If* $U(\theta, (z_0', z_1')) \geq U(\theta, (z_0, z_1))$*, then* $U(\theta', (z_0', z_1')) > U(\theta', (z_0, z_1))$ *for all* $\theta' > \theta$*, and if* $U(\theta, (z_0, z_1)) \geq U(\theta, (z_0', z_1'))$*, then* $U(\theta', (z_0, z_1)) > U(\theta', (z_0', z_1'))$ *for all* $\theta' < \theta$*.*

Proof. Given our assumptions, we have $U(\theta, (z_0', z_1')) \geq U(\theta, (z_0, z_1))$ if and only if

$$u(\omega_0, z_0') - u(\omega_0, z_0) \geq \frac{\theta}{1-\theta}[u(\omega_1, z_1) - u(\omega_1, z_1')].$$

For $\theta' > \theta$, we have

$$\frac{\theta}{1-\theta}[u(\omega_1, z_1) - u(\omega_1, z_1')] > \frac{\theta'}{1-\theta'}[u(\omega_1, z_1) - u(\omega_1, z_1')],$$

in which the inequality follows because $u(\omega_1, z_1) < u(\omega_1, z_1')$ for $z_1 < z_1'$. Together, these inequalities imply $u(\omega_0, z_0') - u(\omega_0, z_0) > \frac{\theta'}{1-\theta'}[u(\omega_1, z_1) - u(\omega_1, z_1')]$, which implies $U(\theta', (z_0', z_1')) > U(\theta', (z_0, z_1))$. The proof of the second claim is analogous. ■

Proof of Proposition 4

We now construct the new regulatory policy $\hat{\mathcal{R}}$ to implement $\hat{\mathbf{a}}$ using the preceding lemmas to ensure that the distance between the $\mathcal{R}$ and $\hat{\mathcal{R}}$ is Lipschitz continuous in the distance between $\mathbf{a}$ and $\hat{\mathbf{a}}$. This means the new policy is within the desired bound for some sufficiently large K.

Proof. Let $\epsilon = d(\hat{\mathbf{a}}, \mathbf{a}) \leq \overline{d} := \max_{\mathbf{a}',\mathbf{a}''\in\mathcal{C}} d(\mathbf{a}', \mathbf{a}'')$. For each signal s, we can divide the set of contracts in M^s into three types: those selected on path (call this set $\mathcal{N}_s^o$), those not selected on path but having higher transfers in ω_1 as compared to any contract in $\mathcal{N}_s^o$ (call this set $\mathcal{N}_s^+$) and those not selected on path but having higher transfers in ω_0 than any contract in $\mathcal{N}_s^o$ (call this set $\mathcal{N}_s^-$).[2] By the assumption that $\mathcal{R}$ satisfies the option diversity condition, each contract in $\mathcal{N}_s^+$ must not underinsure and each contract in $\mathcal{N}_s^-$ must not overinsure.

We will construct, for each s, sets of contracts $\widetilde{\mathcal{N}}_s^+$ and $\widetilde{\mathcal{N}}_s^-$ containing modified versions of contracts in $\mathcal{N}_s^+$ and $\mathcal{N}_s^-$ respectively and at most two new latent contracts. Each type who would choose a contract in $\mathcal{N}_s^+ \cup \mathcal{N}_s^-$ from M^s will continue to purchase the modified version of that contract when offered $\hat{M}^s = M_{\hat{\mathbf{a}}}^s \cup \widetilde{\mathcal{N}}_s^+ \cup \widetilde{\mathcal{N}}_s^-$,[3] their utility increases from the modified contract relative to the unmodified contract, and the firm's profit from that type buying the modified contract is lower than purchasing the unmodified contract. Moreover, any type not in the support of s that chooses the contract in $M_{\mathbf{a}}^s$ with the largest payment in ω_1 will now choose the new downward latent contract from $\hat{M}^s$, and any type not in the support of s that chooses the contract in $M_{\mathbf{a}}^s$ with the largest payment in ω_0 will now choose the new upward latent contract from $\hat{M}^s$.

Let $\mathcal{N}_s^+ = \{(z_0^{n,+}, z_1^{n,+}) : n = 1, ..., N_s^+\}$, in which $N_s^+ = |\mathcal{N}_s^+|$, with indices ordered so that $z_0^{n,+} > z_0^{n+1,+}$ and $z_1^{n,+} < z_1^{n+1,+}$. Similarly, define $\mathcal{N}_s^- = \{(z_0^{n,-}, z_1^{n,-}) : n = 1, ..., N_s^-\}$ with $N_s^- = |\mathcal{N}_s^-|$ and indices ordered so that $z_0^{n,-} > z_0^{n+1,-}$ and $z_1^{n,-} < z_1^{n+1,-}$. Define $\mathcal{N}_s^o$ and N_s^o analogously. To simplify notation, we write $(z_0^{0,+}, z_1^{0,+}) := (z_0^{N_s^0,o}, z_1^{N_s^o,o})$ for the contract in $\mathcal{N}_s^o$ with the largest payment in ω_1 and $(z_0^{0,-}, z_1^{0,-}) := (z_0^{1,o}, z_1^{1,o})$ for the contract in $\mathcal{N}_s^o$ with the largest payment in ω_0. We proceed in steps.

Step 1: Construct new latent contracts

Suppose $\Theta^s(z_0^{0,+}, z_1^{0,+}) \neq \emptyset$ or $\Theta^s(z_0^{0,-}, z_1^{0,-}) \neq \emptyset$; otherwise, continue to the next step. Consider the case when $\Theta^s(z_0^{0,+}, z_1^{0,+}) \neq \emptyset$. Take $\theta_0^+ := \min\{\theta > \overline{\theta}^s\}$. The monotonicity of demand for insurance in θ implies that if any type $\theta > \overline{\theta}^s$ is purchasing $(z_0^{0,+}, z_1^{0,+})$, θ_0^+ must be doing so. Option diversity then implies that $(z_0^{0,+}, z_1^{0,+})$ does not underinsure. Using Lemma 1, construct a new downward latent contract as $(\hat{z}_0^{0,+}, \hat{z}_1^{0,+}) = (z_0^{0,+} - \Delta((z_0^{0,+}, z_1^{0,+}), \alpha_0^+, \theta_+^0), z_1^{0,+} + \alpha_0^+)$, where we increase the net transfer in ω_1 by α_0^+ and decrease the net transfer in ω_0 by $\Delta((z_0^{0,+}, z_1^{0,+}), \alpha_0^+, \theta_+^0)$. We take α_0^+ to be the smallest value of $\alpha \geq \frac{4\epsilon}{\Delta_\Theta \underline{\mu}}$ such that

$$U(\theta_0^+, (z_0^{0,+} - \Delta((z_0^{0,+}, z_1^{0,+}), \alpha, \theta_0^+), z_1^{0,+} + \alpha)) - U(\theta_0^+, (z_0^{0,+}, z_1^{0,+})) \geq \epsilon \overline{u}_z,$$
$$U(\overline{\theta}^s, (z_0^{0,+} - \Delta((z_0^{0,+}, z_1^{0,+}), \alpha, \theta_0^+), z_1^{0,+} + \alpha)) - U(\overline{\theta}^s, (z_0^{0,+}, z_1^{0,+})) \leq -\epsilon \overline{u}_z,$$

so $\theta > \overline{\theta}^s$ prefers the newly constructed latent contract $(\hat{z}_0^{0,+}, \hat{z}_1^{0,+})$ to $(z_0^{0,+}, z_1^{0,+})$, while $\overline{\theta}^s$ prefers the latter.

[2]Note that if a contract $c \in M^s$ is not chosen by any type in *supp*(s) but is chosen by some type $\theta \notin$ *supp*(s), it must be that $\theta > \overline{\theta}^s$ or $\theta < \underline{\theta}^s$ due to the assumption that *supp*$(s) = \{\theta \in \Theta | \underline{\theta}^s \leq \theta \leq \overline{\theta}^s\}$. If the former, c must have a higher net transfer in ω_1 than any contract in M^s chosen by types in *supp*(s), and if the latter, it must have a lower transfer in ω_0.

[3]Recall that $M_{\hat{\mathbf{a}}}^s$ denotes the set of contracts allocated to types in the support of s under $\hat{a}$.

Because the above inequalities hold for α such that $f(\alpha) \geq \epsilon\overline{u}_z$ (where f is as defined in Lemma 1), we have $\alpha_0^+ \leq \max\{f^{-1}(\epsilon\overline{u}_z), \frac{4\epsilon}{\underline{\Delta_\Theta \mu}}\}$. Let $\delta_0^+(\epsilon) = \max\{f^{-1}(\epsilon\overline{u}_z), \frac{4\epsilon}{\underline{\Delta_\Theta \mu}}\}$. Because f is differentiable and strictly increase, f^{-1} is differentiable and therefore Lipschitz continuous on $[0, \overline{d}]$, which immediately implies that $\delta_0^+(\epsilon)$ is Lipschitiz continuous in ϵ on $[0, \overline{d}]$. Moreover, the distance between $(\hat{z}_0^{0,+}, \hat{z}_1^{0,+})$ from $(z_0^{0,+}, z_1^{0,+})$ is $\alpha_0^+ + \Delta((z_0^{0,+}, z_1^{0,+}), \alpha_0^+, \theta) \leq \alpha_0^+(1 + \frac{\theta_0^+}{1-\theta_0^+}) \leq \delta_0^+(\epsilon)(1 + \frac{\theta_0^+}{1-\theta_0^+})$, which is Lipschitz in ϵ.

Take any $\theta > \overline{\theta}^s$. We now show that θ will prefer the new downward latent contract $(\hat{z}_0^{0,+}, \hat{z}_1^{0,+})$ to any contract $(z_0', z_1') \in M_{\hat{\mathbf{a}}}^s$. For an arbitrary $(z_0', z_1') \in M_{\hat{\mathbf{a}}}^s$, let (z_0, z_1) be a contract in $M_{\mathbf{a}}^s$ such that $||(z_0', z_1') - (z_0, z_1)|| \leq \epsilon$. We then have

$$\begin{aligned} U(\theta, (\hat{z}_0^{0,+}, \hat{z}_1^{0,+})) - U(\theta, (z_0', z_1')) &\geq U(\theta, (z_0^{0,+}, z_1^{0,+})) + \epsilon\overline{u}_z - U(\theta, (z_0', z_1')) \\ &\geq U(\theta, (z_0^{0,+}, z_1^{0,+})) + \epsilon\overline{u}_z - U(\theta, (z_0, z_1)) - \epsilon\overline{u}_z \\ &\geq 0, \end{aligned}$$

in which the first inequality follows by our selection of α_0^+, the second inequality by Lemma 3 and last inequality follows because $\overline{\theta}^s$ prefers $(z_0^{0,+}, z_1^{0,+})$ to (z_0, z_1) as agent incentive compatibility holds for $\mathbf{a}$, so every type $\theta > \overline{\theta}^s$ also prefers $(z_0^{0,+}, z_1^{0,+})$ to (z_0, z_1). A similar argument implies no type $\theta \leq \overline{\theta}^s$ prefers $(\hat{z}_0^{0,+}, \hat{z}_1^{0,+})$ to their preferred contract in $M_{\hat{\mathbf{a}}}^s$.

We can similarly construct a new upward latent contract $(\hat{z}_0^{0,-}, \hat{z}_1^{0,-})$ when $\Theta^s(z_0^{0,-}, z_1^{0,-}) \neq \emptyset$ using Lemma 2, with analogous properties.

Step 2: Modify existing off-path contracts

Suppose $\mathcal{N}_s^+ \neq \emptyset$. For $n = 1, ..., N_s^+$, let $\theta_n^+ = \min\{\theta' \in \Theta^s(z_0^{n,+}, z_1^{n,+})\}$. If $\Theta^s(z_0^{0,+}, z_1^{0,+}) = \emptyset$, let $\delta_0^+(\epsilon) = \frac{4\epsilon}{\underline{\Delta_\Theta \mu}}$. Again using Lemma 1, define $(\hat{z}_0^{n,+}, \hat{z}_1^{n,+}) = (z_0^{n,+} - \Delta((z_0^{n,+}, z_1^{n,+}), \alpha_n^+, \theta_n^+), z_1^{n,+} + \alpha_n^+)$ in which α_n^+ is the smallest value of $\alpha \geq \delta_{n-1}^+(\epsilon)$ such that, for $\theta_n' = \max\{\theta \in \Theta : \theta < \theta_n^+\}$, we have

$$\begin{aligned} &U(\theta_n^+, (z_0^{n,+} - \Delta((z_0^{n,+}, z_1^{n,+}), \alpha, \theta_n^+), z_1^{n,+} + \alpha)) - U(\theta_n^+, (\hat{z}_0^{n-1,+}, \hat{z}^{n-1})) \geq k\delta_{n-1}^+(\epsilon), \\ &U(\theta_n', (z_0^{n,+} - \Delta((z_0^{n,+}, z_1^{n,+}), \alpha, \theta_n^+), z_1^{n,+} + \alpha)) - U(\theta_n', (\hat{z}_0^{n-1,+}, \hat{z}^{n-1})) \leq -k\delta_{n-1}^+(\epsilon) \end{aligned}$$

where k is the constant from Lemma 1. Thus, $\alpha_n^+ \leq \delta_n^+(\epsilon) := \max\{f^{-1}(k\delta_{n-1}^+(\epsilon)), \delta_{n-1}^+(\epsilon)\}$. To show that $\delta_n^+(\epsilon)$ is Lipschitz in ϵ for $\epsilon \in [0, \overline{d}]$, we proceed by induction. We have the base case δ_0^+, and suppose $\delta_{n-1}^+(\epsilon)$ is Lipschitz. Because f is differentiable, $f^{-1}(k\delta_{n-1}^+(\epsilon))$ is Lipschitz in ϵ for $\epsilon \in [0, \overline{d}]$, so $\delta_n^+(\epsilon)$ is clearly Lipschitz in ϵ on this interval as well. Because $\Delta((z_0^{n,+}, z_1^{n,+}), \alpha, \theta_n^+) \in [0, \frac{\theta_n^+}{1-\theta_n^+}\alpha]$ by Lemma 1, we conclude that the distance between $(\hat{z}_0^{n,+}, \hat{z}_0^{n,+})$ and $(z_0^{n,+}, z_1^{n,+})$ is at most $\alpha_n^+ + \Delta((z_0^{n,+}, z_1^{n,+}), \alpha_n^+, \theta_n^+) \leq \delta_n^+(\epsilon)(1 + \frac{\theta_n^+}{1-\theta_n^+})$, which is Lipschitz in ϵ for $\epsilon \in [0, \overline{d}]$.

We note that this construction implies $\alpha_n^+ \geq \alpha_{n-1}^+$. Using the same argument as in step 1, we can show that all $\theta \geq \theta_n^+$ prefer $(\hat{z}_0^{n,+}, \hat{z}_1^{n,+})$ to any contract in $M_{\hat{\mathbf{a}}}^s \cup \{(\hat{z}_0^{m,+}, \hat{z}_1^{m,+})\}_{m=0}^{n-1}$

(note that this contains the new downward latent contract) and all $\theta < \theta_n^+$ prefer some contract in $M_{\hat{\mathbf{a}}}^s \cup \{(\hat{z}_0^{m,+}, \hat{z}_1^{m,+})\}_{m=0}^{n-1}$ to $(\hat{z}_0^{n,+}, \hat{z}_1^{n,+})$.

We can construct a similar set of modified contracts $\{(\hat{z}_0^{n,-}, \hat{z}_1^{n,-})\}_{n=1}^{N_s^-}$ for each contract in $\mathcal{N}_0^-$ to go with the new downward latent contract—these contracts have higher net transfers in ω_0 and lower transfers in ω_1, and we define α_n^- analogously.

Step 3: Define the new regulatory policy $\hat{\mathcal{R}}$ and show it implements $\hat{\mathbf{a}}$

Let $\hat{M}^s = M_{\hat{\mathbf{a}}}^s \cup \{(\hat{z}_0^{n,-}, \hat{z}_1^{n,-})\}_{n=0}^{N_s^-} \cup \{(\hat{z}_0^{n,+}, \hat{z}_1^{n,+}\}_{n=0}^{N_s^+}$. Denote by $(\hat{z}_0, \hat{z}_1)_\theta^s$ and $(z_0, z_1)_\theta^s$ the contracts received by type θ with signal s from the allocations $\hat{\mathbf{a}}$ and $\mathbf{a}$ respectively. By construction, we have

$$
\begin{aligned}
&(\hat{z}_0^{n,+}, \hat{z}_1^{n,+}) \in \underset{(z_0', z_1') \in \hat{M}^s}{\arg\max}\ U(\theta, (z_0', z_1'))\ \forall \theta \in \Theta^s(z_0^{n,+}, z_1^{n,+}),\ n = 0, ..., N_s^+, \\
&(\hat{z}_0^{n,-}, \hat{z}_1^{n,-}) \in \underset{(z_0', z_1') \in \hat{M}^s}{\arg\max}\ U(\theta, (z_0', z_1'))\ \forall \theta \in \Theta^s(z_0^{n,-}, z_1^{n,-}),\ n = 0, ..., N_s^+, \\
&(\hat{z}_0, \hat{z}_1)_\theta^s \in \underset{(z_0', z_1') \in \hat{M}^s}{\arg\max}\ U(\theta, (z_0', z_1'))\ \forall \theta \in supp(s).
\end{aligned}
$$

The first two lines imply that it is incentive compatible for $\theta > \overline{\theta}^s$ who choose $(z_0^{n,+}, z_1^{n,+})$ from M^s in the implementation of $\mathbf{a}$, to choose $(\hat{z}_0^{n,+}, \hat{z}_1^{n,+})$ from $\hat{M}^s$ and for $\theta < \underline{\theta}_s$ who choose $(z_0^{n,-}, z_1^{n,-})$ from M^s in the implementation of $\mathbf{a}$ to choose $(\hat{z}_0^{n,-}, \hat{z}_1^{n,-})$ from $\hat{M}^s$. The third line implies that, when $\hat{M}^s$ is offered after signal s, each $\theta \in supp(s)$ finds it incentive compatible to purchase the contract they are specified to receive in $\hat{\mathbf{a}}$. We assume below that types in $s' \neq s$ follow such strategies when offered $\hat{M}^s$.

Because $||(z_0', z_1')_\theta^{s'} - (z_0, z_1)_\theta^{s'}|| \leq \epsilon$, the firm's expected profits in $\hat{\mathbf{a}}$ after s' is at most ϵ less than after s' under $\mathbf{a}$:

$$
\sum_{\theta \in supp(s')} \mu(\theta|s')\Pi(\theta, (\hat{z}_0, \hat{z}_1)_\theta^{s'}) \geq -\epsilon + \sum_{\theta \in supp(s')} \mu(\theta|s')\Pi(\theta, (z_0, z_1)_\theta^{s'}). \tag{3}
$$

The order condition implies that for $s' > s$, there are types in the support of s', namely $\theta > \overline{\theta}^s$, that will find it optimal to purchase a contract from $\{z_0^{n,+}, z_1^{n,+}\}_{n=0}^{N_s^+}$—namely, $\Theta^s(z_0^{n,+}, z_1^{n,+}) \neq \emptyset$ for some $n \in \{0, ..., N_s^+\}$. The probability of such types is bounded below by $\underline{\mu}$. The firm's expected profits from offering $\hat{M}^s$ after signal $s' > s$ when $\theta > \overline{\theta}_s$

make these choices is

$$\sum_{\theta \in supp(s)} \mu(\theta|s')\Pi(\theta, (\hat{z}_0, \hat{z}_1)^s_\theta) + \sum_{n=0}^{N_s^+} \sum_{\theta \in \Theta^s(z_0^{n,+}, z_1^{n,+})} \mu(\theta|s')\Pi(\theta, (\hat{z}_0^{n,+}, \hat{z}_1^{n,+})) \tag{4}$$

$$\leq \sum_{\theta \in supp(s)} \mu(\theta|s')(\Pi(\theta, (z_0, z_1)^s_\theta) + \epsilon) + \sum_{n=0}^{N_s^+} [\sum_{\theta \in \Theta^s(z_0^{n,+}, z_1^{n,+})} \mu(\theta|s')(\Pi(\theta, (z_0^{n,+}, z_1^{n,+})) - \frac{\underline{\Delta}_\Theta}{2}\alpha_n^+)]$$

$$\leq \epsilon - \frac{\underline{\mu}\underline{\Delta}_\Theta}{2} \frac{4\epsilon}{\underline{\Delta}_\Theta \underline{\mu}} + \sum_{\theta \in supp(s)} \mu(\theta|s')\Pi(\theta, (z_0, z_1)^s_\theta) + \sum_{n=0}^{N_s^+} \sum_{\theta \in \Theta^s(z_0^{n,+}, z_1^{n,+})} \mu(\theta|s')\Pi(\theta, (z_0^{n,+}, z_1^{n,+}))$$

$$= -\epsilon + \sum_{\theta \in supp(s)} \mu(\theta|s')\Pi(\theta, (z_0, z_1)^s_\theta) + \sum_{n=0}^{N_s^+} \sum_{\theta \in \Theta^s(z_0^{n,+}, z_1^{n,+})} \mu(\theta|s')\Pi(\theta, (z_0^{n,+}, z_1^{n,+}))$$

$$= -\epsilon + \overline{\Pi}(M^s, s'),$$

where the last line is ϵ less than the profit for the firm from offering M^s after s', which we call $\overline{\Pi}(M^s, s')$. The first inequality above follows from Lemma 1 and $||(z_0', z_1')^{s'}_\theta - (z_0, z_1)^{s'}_\theta|| \leq \epsilon$, the second inequality since $\alpha_n^+ \geq \frac{4\epsilon}{\underline{\Delta}_\Theta \underline{\mu}}$ for all n. Thus, the modified latent contracts in $\hat{M}^s$ ensure the firm's profits from deviating to s' go down relative to the firm's profits from the same deviation under $\mathcal{R}$.

Because $\mathcal{R}$ implements $\mathbf{a}$, the firm's expected profits from offering M^s after signal s' must be less than its profits of offering $M^{s'}$, namely

$$\overline{\Pi}(M^s, s') \leq \sum_{\theta \in supp(s')} \mu(\theta|s')\Pi(\theta, (z_0, z_1)^{s'}_\theta).$$

This inequality along with (3) implies that $\overline{\Pi}(M^s, s')$ is at least ϵ less than the profit from offering $M^{s'}$ after s'. By (4), we know that the profit from offering $\hat{M}^{s'}$ after s' is at most ϵ less than the profit from offering $M^{s'}$ after s'. Putting these together, we conclude that the firm's profit from offering $\hat{M}^s$ after s' is

$$\sum_{\theta \in supp(s)} \mu(\theta|s')\Pi(\theta, (\hat{z}_0, \hat{z}_1)^s_\theta) + \sum_{n=0}^{N_s^+} \sum_{\theta \in \Theta^s(z_0^{n,+}, z_1^{n,+})} \mu(\theta|s')\Pi(\theta, (\hat{z}_0^{n,+}, \hat{z}_1^{n,+}))$$

$$\leq \sum_{\theta \in supp(s')} \mu(\theta|s')\Pi(\theta, (\hat{z}_0, \hat{z}_1)^{s'}_\theta),$$

where the last term is the firm's expected profit after signal s' under allocation $\hat{\mathbf{a}}$. Thus, offering $\hat{M}^s$ after s' gives lower expected profit than offering $\hat{M}^{s'}$. A similar argument applies for $s' < s$.

Constructing analogous $\hat{M}^s$ for each signal s, we define $\hat{\mathcal{R}} = \{\hat{M}^s\}_{s \in S}$. We claim that $\hat{\mathcal{R}}$ implements $\hat{\mathbf{a}}$. We construct an equilibrium in which the firm offers $\hat{M}^s$ after signal s and each type $\theta \in supp(s)$ chooses $(\hat{z}_0, \hat{z}_1)^s_\theta$ from $\hat{M}^s$. If $\theta > \max\{\theta \in supp(s)\}$, then $\theta \in \Theta^s(z_0^{n,+}, z_1^{n,+})$ for some $n = 0, ..., N_s^+$, in which case we specify that θ, when offered $\hat{M}^s$ chooses $(\hat{z}_0^{n,+}, \hat{z}_1^{n,+})$. Similarly, if $\theta < \min\{\theta \in supp(s)\}$, then $\theta \in \Theta^s(z_0^{n,-}, z_1^{n,-})$ for some $n = 0, ..., N_s^-$, in which case we specify that θ, when offered $\hat{M}^s$ chooses $(\hat{z}_0^{n,-}, \hat{z}_1^{n,-})$. As argued above, this is incentive compatible for agents by construction of our latent contracts and the fact that $\hat{\mathbf{a}}$ is agent-incentive compatible. As we have shown above, the firm would have no profitable deviations from offering menu $\hat{M}^s$ after signal s, so firm incentives are satisfied, and $\hat{\mathcal{R}}$ implements $\hat{\mathbf{a}}$. By construction, $\hat{\mathcal{R}}$ uses contracts whose distance from contracts in $\mathcal{R}$ are Lipschitz in ϵ and statement of the Proposition follows. Moreover, by our construction, each $\theta > \overline{\theta}^s$ purchases a downward latent contract that does not underinsure and each $\theta < \underline{\theta}_s$ purchases an upward latent contract that does not overinsure. Thus, $\hat{\mathcal{R}}$ satisfies the option diversity condition. ■

Proof of Proposition 5

Proof. We make the construction using only data identified through the inferences described in section 4.2.

Since M_i is a profit-maximizing contract after signal s_i, the firm has no incentive to offer any other menu after signal s_i. Therefore, if we add latent contracts to M_i which are not chosen by any type in the support of s_i, the only deviations we need to worry about are the firm offering some M_i for $i < n$ after signal s_n. As the regulator knows the maximal risk types $\overline{\theta}^{s_i}$ in the support of s_i for each i, she constructs the latent contracts targeted at this type. For each $i < n$, $\overline{c}^i = (\overline{p}^i, \overline{t}^i)$ is constructed so that no type in the support of signal s_i strictly prefers it from $M_i' = M_i \cup \{\overline{c}^i\}$, while any type $\theta > \overline{\theta}^{s_i}$ prefers it from M_i'. Therefore, if the firm offers M_i' after signal s_n, we can construct equilibria in which types $\theta > \overline{\theta}^{s_i}$ select $\overline{c}^i$, while all other types in the support of s_n select a contract from M_i. If the firm offers M_n^C after signal s_n, every type chooses either (p^C, ℓ) or the contract they would chose from M_n in the laissez-faire outcome.[4]

We will construct $\overline{c}^i(\delta)$ as a continuous function of $\delta := p_n^{m_n} - p^C$, setting $\overline{c}^i(\delta) = c_i^{m_i}$ for $\delta \leq 0$.[5] For $p^C = p_n^{m_n}$ the price cap is non-binding. As we start to lower p^C further, we can continuously bound the change in the firm's profit from offering M_n^C after s_n, as a function of δ. As we lower p^C, some types may switch to purchasing (p^C, ℓ) from M_n^C. By Lemma 1 of Chade and Schlee [2012], these switches must weakly increase the firm's profit: that is, at a price cap p^C such that a type is indifferent between a contract (p, t) and (p^C, ℓ), with

[4]Note that if the price cap is binding for any contract, then it must bind for the highest coverage contract on the menu. For any type θ who chose (p, t) from menu M_i in the laissez-faire outcome, they still prefer (p, t) over any contract that is unaffected by the price cap. Moreover, if the price cap binds for any contract it has to bind at the top, so (p^C, ℓ) is available and dominates all other contracts with price p^C.

[5]If the price cap is non-binding, we don't need a latent contract, and we set the latent contract equal to the highest coverage contract on M_i, so the distance between the two is zero.

$t < \ell$, the firm's profit from selling that type (p^C, ℓ) is higher than selling (p, t). As these switches can only increase the firm's profit, all decrease in profit must come from the impact of the price cap on menu prices. Therefore lowering the price cap by δ lowers firm profit from offering M_n^C after s_n by at most δ—a continuous bound on the firm's change in profit.

By Lemma 4, types $\theta > \overline{\theta}^{s_i}$ will choose $c_i^{m_i} = (p_i^{m_i}, \ell)$ when offered the menu M_i as $c_i^{m_i}$ is the contract selected by $\overline{\theta}^{s_i}$ and has the highest transfer is ω_1. By Lemma 1,[6] we can continuously modify $\overline{c}^i$ from $c_i^{m_i}$ by increasing the premium and the transfer, such that all types weakly below $\overline{\theta}^{s_i}$ prefer $c_i^{m_i}$ to $\overline{c}^i$, while $\theta > \overline{\theta}^{s_i}$ prefers $\overline{c}^i$. Moreover, the difference between the firm's profits from selling $\overline{c}^i$ and selling $c_i^{m_i}$, to $\theta > \overline{\theta}^{s_i}$, is decreasing continuously in the distance between $\overline{c}^i$ and $c_i^{m_i}$. While the regulator may not know the firm's exact profits from offering M_i after signal s_n, she knows that it is less than the firm's profits from offering M_n after signal s_n (as M_n is a firm-optimal menu after s_n). Because the regulator knows the proportion of $\overline{\theta}^{s_n}$ types in the support of s_n (as $\overline{\theta}^{s_n} > \overline{\theta}^{s_i}$, these types prefer $\overline{c}^i$ from M_i'), and the firm's profit from selling $\overline{c}^i$ to these types, she can calculate a lower-bound on how much firm profits will decrease from offering M_i' after signal s_n relative to offering M_i.[7] The distance between $\overline{c}^i$ and $c_i^{m_i}$ that is needed to reduce firm profits from offering M_i' after s_n by δ is continuous in δ.[8] Therefore for each i, we can construct a latent contract $\overline{c}^i(\delta)$ which is continuous in δ such that the firm offers M_i' after signal s_i, and M_n^C after s_n. ■

2 Extensions

Generalization of the Order Condition

We state Theorem 1 under an order condition that represents a natural monotonicity on risk and coverage needs. This ordering of signals allows the regulator to construct optimal regulatory policy using at most two latent contracts. However, the fact that the regulator can implement any solution to (RP) (and, indeed, any allocation that satisfies firm participation and agent incentives) still holds even under weaker assumptions than Assumption 1. We

[6] Since $t_i^{m_n} = \ell$.

[7] While the regulator may not know what the profits from offering M_i after signal s_n are, she knows that as she modifies $\overline{c}^i$, at least type $\overline{\theta}^{s_n}$ will choose different contracts from M_i and M_i'. Because only types $\theta > \overline{\theta}^{s_i}$ switch from purchasing $c_i^{s_i}$ to $\overline{c}^i$ and all such switches reduce firm profits, she can calculate an upper-bound on how much firm profit's changes in $\overline{c}^i$ by looking at the impact of only type $\overline{\theta}^{s_n}$ switching.

[8] The proof of Lemma 1, now taking the value θ_c as defined in that proof, to be $\overline{\theta}^{s_i}$, shows how to construct a contract $\overline{c}^i$ parameterized by $\alpha \geq 0$ such that $||\overline{c}^i - c_i^{m_i}|| = \alpha + \Delta(c_i^{m_i}, \alpha, \overline{\theta}^{s_n})$ and firm profits changes from any type $\theta \in (\overline{\theta}^{s_i}, \overline{\theta}^{s_n}]$ switching to $\overline{c}^i$ from $c_i^{m_i}$, by $\alpha((1-\theta)\frac{\overline{\theta}^{s_i}}{1-\overline{\theta}^{s_i}} - \theta) < 0$; in expectation, this reduces firm profits from offering M_i' after signal s_n by at least $\alpha|(1-\overline{\theta}^{s_n})\frac{\overline{\theta}^{s_i}}{1-\overline{\theta}^{s_i}} - \overline{\theta}^{s_i}|\lambda(c_n^{m_n})$, where $\lambda(c_n^{m_n})$ is the fraction of type $\overline{\theta}^{s_n}$ agents in the support of signal s_n. If we want to reduce firm profits from offering M_i' to s_n by δ, then we can set $\alpha = \delta(|(1-\overline{\theta}^{s_n})\frac{\overline{\theta}^{s_i}}{1-\overline{\theta}^{s_i}} - \overline{\theta}^{s_i}|\lambda(c_n^{m_n}))^{-1}$ and the distance between the latent contract $\overline{c}^i$ and $c_i^{m_i}$ will be at most $\alpha + \Delta(c_i^{m_i}, \alpha, \overline{\theta}^{s_n}) \leq \delta[(|(1-\overline{\theta}^{s_n})\frac{\overline{\theta}^{s_i}}{1-\overline{\theta}^{s_i}} - \overline{\theta}^{s_i}|\lambda(c_n^{m_n}))^{-1}(1 + \frac{\overline{\theta}^{s_n}}{1-\overline{\theta}^{s_n}})]$ (where the inequality follows from $\Delta(c_i^{m_i}, \alpha, \theta) \leq \frac{\theta}{1-\theta}\alpha$ as shown in Lemma 1).

provide one such assumption below, albeit with a regulatory policy that requires more latent contracts. The assumption relaxes the use of a single state for ordering types, and can also be used to relax the assumption that $u(\cdot, \omega_0, x)$ is the same for all x.

Assumption 1 (General Type Ordering and Signal Monotonicity)**.** *For each pair of signals s and s', there exist $(\theta', x') \in supp(s')$ and $\omega', \omega'' \in \Omega$ such that one of the following two conditions hold:*

(a) $\frac{\theta'(\omega')}{\theta'(\omega'')} > \frac{\theta(\omega')}{\theta(\omega'')}$ *and* $u_z(\cdot, \omega', x') \geq u_z(\cdot, \omega', x)$ *and* $u_z(\cdot, \omega'', x') \leq u_z(\cdot, \omega'', x)$ *for all* $(\theta, x) \in supp(s)$

(b) $\frac{\theta'(\omega')}{\theta'(\omega'')} < \frac{\theta(\omega')}{\theta(\omega'')}$ *and* $u_z(\cdot, \omega', x') \leq u_z(\cdot, \omega', x)$ *and* $u_z(\cdot, \omega'', x') \geq u_z(\cdot, \omega'', x)$ *for all* $(\theta, x) \in supp(s)$

The construction of a regulatory policy to implement $\mathbf{a}$ is analogous to those in the proof of Theorem 1, with ω' playing the role of ω_1 and ω'' playing the role of ω_0. The regulator adds to $M^s_{\mathbf{a}}$ an upward and downward latent contract for each state, constructed in the same way as in Theorem 1 relative to ω_1. They are not purchased by any types in the support of signal s but are attractive to types in signal $s' \neq s$ (which exist in each $s' \neq s$ by Assumption 1) and reduce firm profits by an arbitrarily large amount when purchased. Although we assumed that u is constant across ω_0 in the baseline model, it is easy to see that the proof of Theorem 1 still holds whenever $u_z(z, \omega_0, x)$ is (weakly) decreasing in x.

Observable Agent Characteristics

In some cases, a regulator may be able to condition her policy on observable attributes of individual agents, ensuring that for particular subgroups, the firm can only offer a particular subset of menus. For instance, governments often condition benefits or subsidies on earnings, and employee health plan premiums often vary based on salary—these attributes may well be relevant to agents' preferences over transfers. Our main result applies with essentially no changes in these settings. If the regulator observes a partition of agent categories, she can effectively choose a separate regulatory policy for each element of the partition. As long as the categories within each partition element satisfy our order conditions, Theorem 1 applies within each partition element.

Competition and Market Structure

In many insurance markets, firms face competition. We explore competition within the non-contractible loss model, assuming the regulator wishes to maximize consumer welfare. Due to fixed costs, efficiency requires a single firm in the market. If all firms that enter split the market equally, our analysis extends immediately: the optimal regulatory policy is the same as with one firm, only one firm enters, and this firm receives zero profit—if more firms enter, some firm makes negative expected profit. Though reassuring, equal splitting is not always a reasonable assumption in practice. In a decentralized market, entrants can selectively target agents through advertising. This may facilitate cream-skimming, even if the entrant is legally required to offer insurance to all. Cream-skimming is a serious concern because it 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 agents, and a single potential entrant may target advertisements to particular categories. The entrant must also serve all agents, but we suppose that an agent defaults to the incumbent unless he 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. Assume both firms' signals are perfectly informative of each agent's category, x, and the regulator's objective is to maximize consumer welfare. Moreover, 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 $\kappa + (1 - \alpha)C$.

To illustrate the problem, suppose there are two categories $x \in \{H, L\}$ and two degenerate events $\omega \in \{\omega_0, \omega_1\}$—an agent in category x suffers loss ℓ_x in event ω_1, which occurs with probability θ_x. In the optimal allocation, both types get full insurance at a common price p^*, and the incumbent's participation constraint binds. Selling to all agents never covers the entrant's fixed costs, but targeting only category L, with $\ell_L \leq \ell_H$ 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 is not part of an equilibrium.

How well can the regulator do when entrants might cream-skim? Let $\mathcal{T}_c \subset \mathcal{T}$ denote the set of types 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

$$\pi_e(\mathbf{a}) = \max_{\mathcal{T}_c \subset \mathcal{T}} \sum_{\tau \in \mathcal{T}_c} \alpha_x \mu(\tau)(\Pi(\tau, c_\tau) - C) - \kappa.$$

Suppose fixed costs for the incumbent are large enough that the regulator wants to deter entry (and assume doing so is feasible). Entry is not a problem as long as the allocation satisfies a no cream-skimming constraint which ensures the entrant finds it unprofitable to enter the market:

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

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

The natural analog to the relaxed regulator's problem in this setting is:

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

Assume a solution to this problem exists. The regulator maximizes her objective function subject only to the incumbent firm's participation constraint $\pi(\mathbf{a}) \geq 0$ and the entrant firm's no cream-skimming constraint $\pi_e(\mathbf{a}) \leq 0$. Our relaxed problem in this context assumes the regulator wants to prevent entry. If there are many potential entrants and free entry into the market, then $\pi_e(\mathbf{a}) \leq 0$ must hold in any equilibrium allocation. Therefore, the regulator can do weakly better by deterring entry.

In this modified problem, the regulator targets a single contract to each category, offering full insurance, but the price can differ across categories. Offering lower prices to agents in lower categories is necessary to prevent cream-skimming, but this also limits cross-subsidization.

Proposition 6. *The solution to* (CRP) *provides a contract c_x to all types in category x. There exists a collection of contracts $\{\underline{c}_x, \overline{c}_x\}_{x\in X}$ such that the regulatory policy $\mathcal{R} = \{M_x\}_{x\in X}$ with $M_x = \{\underline{c}_x, c_x, \overline{c}_x\}$ implements the solution to* (CRP).

Proof. The argument that the solution to the relaxed problem provides the same contract to all agents in the same category is analogous to Proposition 1, and we omit it. Implementability is a corollary to the proof of Theorem 1. ∎

Constrained Regulation

Two important policy questions for insurance regulation are whether to prevent firms from excluding customers and whether to mandate buying insurance. While optimal policy requires both features—implementation depends on the firm offering insurance to all agents and all agents buying insurance—regulators may not always have the ability to enforce them. We can capture many natural restrictions on the regulator through constraints on the space of permissible policies. Here are a few examples:

- All agents must have the same options: The policy $\mathcal{R}$ must contain a single menu M. This might reflect anti-discrimination laws or strong fairness norms.
- The regulator can enforce contracts, but not menus: The policy must have the form $\mathcal{R} = 2^{\mathcal{C}} \setminus \{\emptyset\}$ for some set of contracts $\mathcal{C}$. If a regulator cannot verify whether the firm offers specific contracts, then the firm may construct its own menus from some grand set of permitted contracts.

- Firms can refuse service: The policy $\mathcal{R}$ must include the null menu $\{(0,0)\}$. Without a law to ensure otherwise, the firm retains the right to exclude some potential customers.

- Agents can opt out: The policy $\mathcal{R}$ must include the null contract $(0,0)$ in each menu. This corresponds to the inability to enforce a purchase mandate.

Insurance Mandates: A Double-Edged Sword:

Definition 1. *A **regulatory policy with unenforceable menus** is a regulatory policy taking one of two forms. The regulator chooses a set of contracts $\mathcal{C}$ such that either i) $\mathcal{R} = 2^{\mathcal{C}} \setminus \{\emptyset\}$ or ii) $\mathcal{R} = \{S \cup \{(0,0)\} : S \in 2^{\mathcal{C}} \setminus \{\emptyset\}\}$. In the former case, there is an insurance mandate, and in the latter case, there is no insurance mandate.*

We show in a special case of the non-contractible loss model that insurance mandates may harm welfare when the regulator can enforce contracts but not menus. Consider the non-contractible loss model with one loss event, ω_1, and two categories $x \in \{L, H\}$, each with a single risk type. Category L faces risk θ_L of loss ℓ_L in event ω_1, and category H faces risk θ_H of loss ℓ_H. Assume $\ell_H > \ell_L$, initial wealth $e > \ell_L$ and $\theta_H > \theta_L$. A fraction μ of agents are in category L. We assume the regulator seeks to maximize consumer welfare.

Proposition 1 implies that the regulator's optimal solution entails selling full insurance to all categories at the same price, set to give the firm exactly zero profit. To implement this allocation, the regulator must force the firm to include a latent contract in each menu. We study what a regulator can achieve if it can only enforce a set $\mathcal{C}$ of permitted contracts, from which the firm can construct menus as it pleases.

We first explore policies with an insurance mandate. Since there is no need to screen within categories, without loss the firm offers a single contract to each agent, and the universe of permissible contracts is

$$\mathcal{C} = \{(p_h, t_h), (p_L, t_L)\}.$$

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 C to solve

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

The regulator must satisfy the firm's participation constraint and incentive constraints to offer the correct contract to each type.

Under these assumptions, any policy with mandatory insurance optimally has the firm offer the same contract to both categories. The firm IC constraints imply

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

While the regulator wants the firm to offer higher coverage to category H, the best she can do is to ensure both categories get the same contract. If every agent receives the same contract, then at least one category is underinsured or overinsured. Since the firm's participation constraint binds, the regulator's problem amounts to choosing a single coverage level $t \in [\ell_L, \ell_H]$ and setting

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

Proposition 7. *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. 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

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

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

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

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

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

This implies that 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

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

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

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

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

Substituting the second into the first gives

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

Note that

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

which implies $u'(e-p_L) > u'(e - p_H + t_H - \ell_H)$. 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. ∎

This result highlights a problem with insurance mandates under a weak regulator. If the regulator cannot force the firm to include latent contracts in the menus, then the firm can hold agents hostage, offering only the most expensive or lowest coverage option. 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.

Example: Optimality of Optional Insurance

Suppose

$$u(z,\omega,x) = log(100 - p + t(\omega) - \ell_x(\omega)),$$

where initial wealth $e = 100$, and $\ell_x(\omega) = \ell_x$ if $\omega = \omega_1$, and zero otherwise. Assume the two categories L and H are equally prevalent, that $\theta_L = 0.5$ and $\theta_H = 0.501$, and that $\ell_L = 30$ and $\ell_H = 114$. If insurance is mandatory, the optimal policy prescribes a single contract $(p^*, t^*) \approx (47.6, 95.3)$ for all agents. Category L is overinsured, category H is underinsured, and consumer welfare is approximately 4.048. If insurance is optional, the regulator can do better by allowing the contracts $(30, 60)$ and $(55.49, 110.7)$. Category H prefers either contract to the null contract, but the former is unprofitable since $\theta_H > \frac{1}{2}$, so the firm offers $(55.49, 110.7)$ to H. Category L on the other hand would rather go uninsured than pay the high premium. Consumer welfare is approximately 4.09.

In order to simplify the question of *when* a mandate is strictly dominated by no mandate, we focus on the case when there is no fixed cost ($k = 0$) and look at the limit as $\theta_L \to \theta_H$, so that the difference in losses are the main difference between L and H.

Proposition 8. *There exists $\bar{\ell}_H$ such that, if $\ell_H \in [\bar{\ell}_H, \frac{e}{\theta_H})$ and $|\theta_L - \theta_H|$ is sufficiently small, it is optimal to make insurance optional.*

Proof. Consider the optimal contract under the mandate. We know that both types receive (p, t). Because the firm's participation constraint, will bind, we know $p = (\mu\theta_L + (1-\mu)\theta_H)t$. It is straightforward when $\theta_H \approx \theta_L$ to show that $t \in (\ell_L, \ell_H)$. Consider the optimal contract with a mandate when $\ell_H = \frac{e}{\theta_H} - \epsilon$ for some $\epsilon > 0$. In order for H to not receive negative infinite utility, it must be that $e - p + t > \frac{e}{\theta_H} - \epsilon$. Because, when $\theta_L \approx \theta_H$, the optimal payment $p \approx \theta_H t$, we get $e - p < \epsilon \frac{\theta_H}{1-\theta_H}$. As $\epsilon \to 0$, L's utility approaches $-\infty$.

Suppose the regulator drops the mandate and only allows the firm to offer $(p, t) = (\theta_H \ell_H, \ell_H)$. For ℓ_H high enough, the autarky outcome for L is strictly higher utility than under the mandate. H prefers the new contract to the previous mandate contract when $\theta_H \approx \theta_L$ because $t < \ell_H$ and H benefits from moving to a zero-profit contract with underinsurance to a zero-profit contract with full-coverage. ■

Intuitively, the larger the difference in coverage needs between the two categories, the more costly it is to pool them. The optimal policy with an insurance mandate involves overinsurance for category L and underinsurance for category H. As ℓ_H increases, the optimal contract with an insurance mandate converges towards full insurance for category H, which is costly for category L. Without the mandate, when $\ell_L < e$ agents in category L are able to guarantee their autarky payoff, which is strictly higher than the payoff from buying the contract for category H. When the regulator cannot force the firm to offer additional options in particular menus, letting agents opt out provides an alternative way to discipline the firm. This could become more effective if the regulator were to directly offer some insurance contracts—a "government option." Changing the agents' outside option could allow the regulator to maintain more cross-subsidization. The interaction of government options with constrained regulation seems an interesting avenue for further research.

Technical extensions

Many natural extensions are straightforward. For instance, little changes if we assume compact—rather than finite—sets of categories and types. Moreover, one can further extend the implementation result to allow even weaker order conditions on agent categories if one is willing to include additional latent contracts in each menu. Generically, we can strengthen our main result to implement the desired allocation in a way such that some agents have strict incentives to choose latent contracts after firm deviations, and on path, all agents have strict incentives to not choose the latent contracts—to do this, simply adjust the payouts of our latent contracts by ϵ to break indifference and make agent preferences strict. Thus, one can make the firm's menu choice preference strict as well.