The Political Economy of Regulation in Markets with Naïve Consumers

Patrick L. Warren and Daniel H. Wood*

November, 2013

Abstract

In a model of a competitive industry selling base goods and add-ons, we investigate the conditions under which citizen-consumers will support policies that eliminate behavioral inefficiencies induced by naïve consumers. Unregulated competitive markets have two effects: they produce deadweight losses, and they redistribute income away from biased consumers. Both unbiased and naïve consumers believe that they benefit from this redistribution (the naïve consumers are wrong), so support for efficiency-improving regulation is limited. Extending our model to consumers with partial sophistication about their naiveté, we predict patterns of regulation consistent with the form and timing of the Credit Card Accountability Responsibility and Disclosure (CARD) Act of 2009.

JEL Classification: D03, D78

Keywords: naiveté, overconfidence, regulation, market failure, government failure, behavioral economics, behavioral public policy, behavioral industrial organization, Credit CARD Act.

*Department of Economics, Clemson University. patrick.lee.warren@gmail.com, dwood2@clemson.edu. We thank the editors and referees at the Journal, seminar audiences at UC Berkeley, Clemson University, Emory University, the University of South Carolina, and North Carolina State University for helpful questions and comments. We want to specifically thank Daniel Gottlieb, Ryan Bubb, Charles Thomas, and Neil Bhutta.
1 Introduction

Markets with imperfectly rational consumers can exhibit persistent inefficiencies. Many recent papers provide theories and evidence that markets sometimes fail to completely restore efficiency when consumers are imperfectly aware of their biases (Gabaix and Laibson, 2006; Sandroni and Squintani, 2007; Grubb, 2009; Heidhues and Koszegi, 2010; Bubb and Kaufman, 2013). Within the context of these models, appropriate policy responses such as mandatory information provision or restricting contractual forms could increase social welfare. Most behavioral public policy proposes normatively optimal regulations to help biased consumers (Gabaix and Laibson (2006, section IV.B); Bar-Gill (2006, section IV); Heidhues and Koszegi (2010, section II.B)).

Almost no attention has been paid to the positive question of when these policies are likely to be adopted. In this paper, we study the conditions under which a regulator beholden to a set of potentially-biased citizens will choose policies that improve welfare.

In our model, consumers purchase a base good and, probabilistically, an add-on. Consumers need the add-on with either low or high probability, but (in the benchmark model) all consumers believe they will need the add-on with the lower probability. In a perfectly-competitive market for the base good, but with producer-specific add-ons, the na¨ıve implicitly subsidize the rationally low risk. Due to this subsidy, a good with negative net social welfare can be consumed in equilibrium. In these circumstances, many welfare-improving regulations exist, but we show that no incentive-compatible budget-balanced mechanism will both improve social welfare and garner the support of the citizens. Both the naïve and the low-risk believe that they benefit from the implicit subsidy, although only the low-risk are correct in this belief. All welfare-improving mechanisms limit the implicit subsidy, so citizens universally view them as strictly inferior to the market outcome.

There are, nevertheless, sporadic regulations that target inefficient consumption induced by consumer biases. Recent examples are the passage of the Credit Card Accountability Responsibility and Disclosure (CARD) Act in 2009 and the creation

---

1Ellison (2006) surveys the early behavioral IO literature. For a recent synthesis of many of the applications of Behavioral Economics to IO, see Spiegler (2011).

2See Berggren (2012) for a recent quantification of the size of this gap in the literature. He finds over 20 percent of behavioral economics papers in the top 10 journals propose some paternalistic policy action, while fewer than 5 percent of those who propose such an action consider the cognitive abilities of policymakers in the adoption and/or implementation of policy. None of those, to our knowledge, actually analyze the support for these policies among voters.
of the Consumer Financial Protection Bureau (CFPB) in 2010. We investigate the extent to which this pattern of regulation can be explained by consumers being (imperfectly) sophisticated about their naïveté. Specifically, we extend the model to allow consumers to make a regulatory decision “behind the veil” before realizing their risk types and interacting in the market. When making this decision, consumers realize there is some chance that they will be underestimate their risk when acting in the market. If consumers are perfectly sophisticated in this belief (i.e., they think they will be naïve with the exact same probability they will actually be naïve), they will support efficient regulation. But if they are imperfectly sophisticated, they will support regulation in extreme cases only, where the social cost of consumption much outweighs the social benefit.

This extended model also helps to understand when regulation targeted at behavioral weaknesses will be enacted and what form it is likely to take. We illustrate these points in a brief study of the CARD Act. The Act regulated certain kinds of add-ons – surprise-based “gotchas” such as late fees or universal default – but left another kind – high interest rates on persistent indebtedness – largely unconstrained. In our model there is greater support for regulations of contractual terms with low potential for redistribution (small differences between low and high risk, low levels of naïveté, and low maximum add-on price) relative to the level of inefficiency. The two credit-card add-ons differ in exactly the ways that would make the regulation of “gotchas” the more attractive regulatory target in our model.

2 Markets and Regulations with Naïve Consumers—A Benchmark

This section derives our basic results in a standard model of an industry with imperfectly rational consumers and add-ons. Our model is similar to Sandroni and Squintani (2007), who implement naïveté identically, but use their model to analyze the market for insurance.\(^3\) The presence of overconfident agents can reduce the scope for Pareto-improving regulation in a market for insurance with adverse selection. We find a similar effect, in a different market (add-ons), and we relate that fact back to

\(^3\)The Sandroni and Squintani model is most similar to our three-type model in Section 3, of which the benchmark is a special case. Their interest is insurance, so they posit risk-averse consumers and the gains from trade come from the transfer of risk to the risk-neutral firm. Otherwise, the model is identical to the free-market subgame in 3.1, except for notation.
the demand for regulation by analyzing regulation as a mechanism design problem.

It is also similar to Bubb and Kaufman (2013), in that firms are selling add-ons to naïve consumers. They consider the impact of non-profit firms (i.e., credit unions) in a market similar to ours. In their model, no commitment to add-on prices is possible, but charging a large add-on price has non-monetary costs. Non-profits have weaker incentives to take advantage of naïve consumers, so they limit some of the inefficiencies induced by naïveté. The introduction of non-profits leaves scope for regulation, since they do not completely correct the inefficiency.4

2.1 The Benchmark Model

Payoffs and Production Consider a market with a base good and an add-on. There is a unit mass of consumers who receive a utility of $B$ for consuming the base good but may be subject to an event that decreases this value by $k$. We call this event “needing the add-on.” Consumers who do not buy the base good never need the add-on, and those needing the add-on can avoid this utility loss by consuming the add-on.5

A competitive set of firms produce units of the base good at constant marginal cost $C$. The add-on good is costless to produce, and it is specific to the base good produced by any given producer.6 Define $G \equiv B - C$ as the gains from trade in the market.

The game consists of three periods. In the first period, firms set their prices simultaneously, and may offer an array of contracts, where firm $i$ chooses a vector of prices $p_{B,i}$ for the base good and $p_{A,i}$ for the associated add-ons. Consumers also learn their risk type in this period. In the second period, consumers choose whether to buy the base good and from which contract. In the last period, consumers learn whether or not they are subject to the welfare shock and decide whether they want to buy the add-on.

---

4Indeed, the equilibrium outcome in their model, in which high-risk sophisticates patronize non-profits while other types patronize for-profits is akin to the market equilibrium in our model once sophisticated high risk consumers are added (in Section 3). Their ending point is, essentially, our starting point.

5The role of $k$ in our model is to place a limit on the price that sellers can charge for the add-on. An alternative justification is that $k$ measures the cost of some avoidance behavior that the consumer can take.

6The assumption of costless add-on production is made for convenience. The qualitative results do not change if add-on production costs an amount strictly less than $k$. 
Types and Information  There are two types of consumers. Naïve consumers need the add-on with probability $\alpha_h$ and low-risk consumers need it with probability $\alpha_l < \alpha_h$. Naïve consumers make up a fraction $\gamma$ of all consumers, and are certain (incorrectly) that they are low risk. Consumers (believe they) know their own type when selecting among contracts during period two, but firms never observe consumers’ types. The distribution of types is common knowledge, including the presence of the naïve types, as are the firms’ costs. Let $\bar{\alpha} \equiv \gamma \alpha_h + (1 - \gamma) \alpha_l$ represent average add-on risk in the population.

Market Equilibrium

Proposition 1. In every subgame perfect Nash equilibrium of the market game described,

1. If the gains from trade are sufficiently negative ($G < -k(\bar{\alpha} - \alpha_l)$), no trade occurs.

2. Otherwise, the zero-profit contract with maximum add-on price ($p_A^* = k, p_B^* = c - \bar{\alpha}k$) is offered and accepted by all consumers.

Proof. This and other proofs are in the Appendix.

The equilibrium is unique in all essential characteristics. The multiplicities arise from the fact that firms may offer a wide variety of contracts that are never accepted, and consumers will be indifferent among several producers, each offering the same contract.

Competition pushes the add-on price to its maximum level because all consumers think they are low risk. Otherwise an entrant can profitably deviate by offering a contract that exceeds the add-on price of the accepted contract by $\epsilon$ and with a base-good price a bit less than $\alpha_l\epsilon$ above the accepted contract. This deviation will attract customers, since the total expected price is lower, but profits increase, since the firm will collect the add-on price with probability $\bar{\alpha} > \alpha_l$.

There is a wedge between consumers’ subjective expected total price ($C - k(\bar{\alpha} - \alpha_l)$) and the real average total price ($C$). For $G > 0$, this wedge is purely redistributive, but, for modestly negative gains-from-trade ($-k(\bar{\alpha} - \alpha_l) < G < 0$), it has real net welfare effects. Consumers consume goods for which the social cost exceeds the social benefit. This wasteful overconsumption motivates regulation.

7In Section 3 we introduce knowingly high-risk consumers.
We analyze constrained-optimal policy as set by a paternalistic regulator who maximizes a weighted average of the citizens’ objective expected utilities, subject to participation and incentive constraints, and budget balance.\footnote{Without specifying the marginal social value of outside public funds, and how those funds would be distributed among the consumer groups, it is difficult to analyze non-balanced mechanisms. For this reason, we will consider only budget-balanced mechanisms. But the gross benefits of a non-balanced policy to a paternalistic regulator who weights types in accord with their population share is straightforward to analyze. An unbalanced policy allows the regulator to loosen the pseudo-participation constraint (3). When that constraint binds, a paternalistic regulator can take advantage of the relaxation of the constraint to improve social welfare. For large negative gains from trade, he will dictate no consumption, compensating the citizens with a transfer for their subjective loss. For smaller negative gains from trade, the regulator will dictate probabilistic consumption, since the net subsidy is not sufficient to fully compensate for forgone consumption. In the range where the pseudo-participation constraint binds, the regulator’s benefit from a regulation that is unbalanced by a dollar exceeds a dollar by a discrete amount. Thus, if the deadweight cost of tax revenues are second-order, there are economic gains from non-balanced policies.}

To allow the regulator maximum flexibility, we use a mechanism-design approach to regulation, in which citizens are induced to provide information to the regulator about add-on need, which the regulator uses, in turn, to dictate their consumption and payment levels. By the revelation principle, it suffices to consider these truthful, direct mechanisms. A budget-balanced regulatory mechanism consists of a set of messages from the citizens \( M = \{0, 1\} \), where \( M = 1 \) indicates a need for the add-on; potentially message-contingent transfers from each citizen \((t, t^A) \in \mathbb{R}^2\), where \( t \) is a transfer paid by all citizens and \( t^A \) is a transfer paid by those needing the add-on; and potentially message-contingent probabilities of receiving the base good and add-on \((\pi, \pi^A) \in [0, 1]^2\) (where \( \pi^A \) is conditional on \( M = 1 \)). Since \( t \) is always collected, \( t^A \) is collected with probability \( \pi \pi^A \alpha \), and the good is produced with probability \( \pi \), budget balance requires that

\[
t + \pi \pi^A \alpha t^A = \pi C. \tag{1}
\]

Incentive compatibility for those needing the add-on and those not needing it requires that

\[
0 \leq t^A \leq k. \tag{2}
\]

We also require the policy to satisfy a pseudo-participation constraint, representing the idea that regulation is subject to a political constraint where it cannot be implemented over the wishes of the citizens. Thus, consumers should weakly prefer the announced mechanism to the free market, i.e.,

\[
\pi \left[ B - \alpha_l [\pi^A t^A + (1 - \pi^A)k] \right] - t \geq \max \{0, G + (\bar{\alpha} - \alpha_l)k\}. \tag{3}
\]
A paternalistic regulator maximizes the weighted sum of the objective consumer surplus of the two consumer types, with weight $\psi_n$ on the naïve, and weight $\psi_l$ on the low-risk. He solves

$$\max_{t, \pi} \left\{ \psi_n \left[ \pi \left[ B - \alpha_h (\pi^A t^A + (1 - \pi^A) k) \right] - t \right] + \psi_l \left[ \pi \left[ B - \alpha_l (\pi^A t^A + (1 - \pi^A) k) \right] - t \right] \right\},$$

subject to constraints (1-3). The resulting policy is stark.

**Proposition 2.** The only mechanism satisfying (1-3) duplicates the free-market equilibrium.

The market allocation uniquely maximizes the subjective expected welfare of both types, so a regulator restricted to deliver at least that subjective welfare is powerless to alter outcomes. This result is the equivalent of Result 1 in Sandroni and Squiniani (2007), where any compulsory insurance lowers the subjective welfare of those consumers believing themselves to be low risk.

The citizens’ opposition is not as simple as a given consumer being opposed to a regulation that forecloses his ability to choose his free-market contract. He would also be opposed to the (infeasible) regulation that prohibits offering that contract to all naïve consumers, but continued to allow rational consumers to take it up. This opposition arises because, without the naïve consumers participating, the market equilibrium would provide no subsidy to the low-risk. Subsidizing the base good would be a net-loser for firms without the participation of the naïve, and they would, therefore, not offer $p_B < C$.

**Monopoly** A monopolistic producer sets add-on prices exactly like a competitive producer. The monopolist simply collects all the subjective surplus with a larger base-good price. Thus, equilibrium consumption and total welfare are identical to equilibrium consumption in competitive markets.

But an important difference arises in the case of a regulator constrained to deliver the subjectively expected market payoff. The monopolist leaves no subjective consumer surplus to the citizens, so the constraint to deliver at least the free-market payoff (3) becomes the standard participation constraint

$$\pi \left[ B - \alpha_l (\pi^A t^A + (1 - \pi^A) k) \right] - t \geq 0.$$
With this looser constraint, the politically-constrained regulator can now use regulation to increase total welfare.

**Proposition 3.** Consider the case of a paternalistic regulator subject to constraints (1, 2, and 5).

1. If the regulator is weakly biased in favor of the naïve \((\psi_n \geq \gamma)\), any constrained-optimal regulatory mechanism dictates universal consumption if and only if there are positive gains from trade (i.e., \(\pi = \pi^A = 1\) if \(G > 0\) and \(\pi = 0\) if \(G < 0\)). Whenever there is consumption and the regulator’s bias is strict, there is no additional transfer required for the add-on \((t^A = 0\) if \(\psi_n > \gamma\)).

2. If the regulator is biased in favor of the rational \((\psi_n < \gamma)\), the optimal regulatory mechanism dictates universal consumption if and only if gains from trade are above a negative threshold (i.e., \(\pi = \pi^A = 1\) if \(G > -k[\alpha - (\psi_n \alpha_h + \psi_l \alpha_l)]\), else \(\pi = 0\)). Whenever there is consumption, the pricing duplicates the competitive free market \((t^A = k)\).

A regulator can only ever redistribute from the naïve to the low risk, because neither group can reliably self-identify ex ante and a rebate to those claiming to need the add-on is never incentive compatible. If he is unbiased or biased in favor of the naïve, he has no desire to enact this redistribution, and so his best policy is to charge everyone \(C\) and dictate consumption whenever \(G > 0\).

When the regulator is biased in favor of the rational, he can enact a regulation that mirrors the cross-subsidization that occurs in the competitive market. For \(G > 0\), the drive to redistribute has no effect on the overall surplus to be divided, but if \(G < 0\), there is a trade-off from the regulator’s perspective. Allowing consumption reduces overall welfare, but also allows for redistribution from the naïve to the low risk, and he balances these two considerations.

The pro-rational case may be particularly relevant if the regulator is subject to lobbying by the citizens, since all the naïve actually believe they are low risk. If the regulator, for instance, is induced to put even a small weight on subjective welfare, as he might in a Grossman and Helpman (1994) lobbying model where some or all citizens are organized, he would end up maximizing an implicitly pro-rational welfare function. Such a regulator, facing a monopoly, would shut down only those market with the most extreme negative gains from trade.
2.2 Discussion

The key conclusions from our benchmark model are: 1) inefficient overconsumption may arise in free markets and 2) a welfare-maximizing regulator may fail to implement feasible regulation to correct this overconsumption if he is either a) required to deliver at least the subjective free-market payoff or b) induced to put some weight on the subjective payoff of consumers.

For the range of negative $G$’s $(-k\gamma(\alpha_h - \alpha_l) < G < 0)$, naïveté leads to inefficient overconsumption. This range increases in three factors: the disutility of needing the add-on and not receiving it, $k$; the share of naïve consumers, $\gamma$; and the risk differential $\alpha_h - \alpha_l$. Furthermore, if the regulator has any scope for action but favors the rational, the range for which he will not intervene to fix the problem (part 2 of Proposition 3) also increases in those factors, as well as in the weight he places on the low-risk in his objective function ($\psi_l$). Taking the model to the world, we expect to find overconsumption existing in equilibrium and failing to be corrected by regulation when the transfer inducing the overconsumption is large, there are many who misperceive their risks, and the regulator is either restricted to interventions that everyone would accept over the free market or he is induced to underweight the behavioral types in his social welfare function, perhaps due to lobbying.

Loosely speaking, markets with naïve consumers lead to too little demand for efficient regulation. Several other papers investigate how biases might distort the political or regulatory “markets”, but they focus mainly on the behavior of politicians. Glaeser (2006) is pessimistic about the chances of successful behavioral public policy because decision makers will often be even more prone to bias than private decision makers. For instance, consumers have weaker incentives to choose well when voting for politicians than when buying products or services and individual politicians may be less costly to mislead than a mass of consumers. Similarly Bisin et al. (2013) show that the political process undermines attempts of sophisticated time inconsistent citizens to commit to their long-term selves’ preferred consumption profiles, with the collective action problem of government debt and private citizens’ problem of time-inconsistency interacting to the social detriment. Finally, Lizzeri and Yariv (2013) analyze how politicians responding to time-inconsistent citizens will provide costly commitment technology and dictate the timing of consumption, if they are delegated these policy levers. The key result is that either full delegation of both levers is optimal or no delegation is optimal, depending to the severity of the time-inconsistency.

Many assumptions in our benchmark model are stark. First, citizens are per-
fectly rational and self-interested other than having potentially biased beliefs about $\alpha$. Citizens evaluate regulatory proposals using exactly the same criteria they do for the market proposals. In reality, however, regulation is not usually presented as a complete direct mechanism, which would exactly parallel the price-based information available under the market, and evaluating a proposal may require anticipating equilibrium responses. For example, a regulation may impose price caps on add-ons without directly speaking to base-good prices (as the CARD Act does). Not all consumers need to calculate the equilibrium effects, however, if there are industry groups, reporters, or other sources of information to help them. Nearly all newspaper articles about the CARD Act mentioned that it could lead to increased fees and reduced rewards. That said, a citizen who underestimates (overestimates) the increase in base-good prices would support a broader (narrower) set of regulations. Second, altruistic citizens might support regulation not in their self-interest. We analyze altruism formally in Section 3.3.

In our model, the policy instrument is, essentially, a tax-and-transfer regime. Other policy responses are available. Governments could attempt to “debias” consumers, perhaps by requiring firms to educate their customers. Bertrand and Morse (2011) find that consumer information regulations might be an effective policy tool when it comes to payday borrowing. But many choices are complex enough that relevant information cannot be quickly summarized (Hastings and Weinstein, 2008; Mastrobuoni and Weinberg, 2009). Another policy response frequently proposed by behavioral economists is light, “libertarian” regulation, such as a “nudge” towards certain options via strategically setting a default option (Camerer et al., 2003; Thaler and Sunstein, 2009). While we model regulation as a general budget-balanced mechanism, “softer” regulatory interventions would fall prey to the same under-demand problems that we identify if they induced similar consumption and transfers to the optimal mechanism.9

3 Regulation with Partially Sophisticated Consumers

In the benchmark model, citizens oppose regulation because they believe that they are benefiting from others’ naïveté. In fact the Consumer Financial Protection Bureau (a product of the Dodd-Frank Act) and the 2009 Credit Card Accountability and Dis-

9Obviously heavy-handed regulations can cause greater distortions than our efficient mechanisms would, but we abstract away from these implementation problems. If voters are also skeptical about the efficiency of government regulation, regulation is even more unlikely to occur in equilibrium.
closure (CARD) Act both attempt to regulate financial products that are appealing in part because of consumer biases. While the post-Lehmann, post-bailout zeitgeist may have been the proximate trigger, a variant of our model in which consumer’s make regulatory decisions at some distance from their market transactions can help explain why the push for regulation was successful.

We alter the benchmark model in two ways. First, we introduce a third type of consumer, the knowingly high-risk. Let $\beta$ be the fraction of all consumers who have $\alpha = \alpha_h$ and now interpret $\gamma$ as the fraction of these consumers who believe themselves to have $\alpha = \alpha_l$. There are now three types: “high-risk” consumers (who are fraction $\beta(1 - \gamma)$ of all consumers), “low-risk” consumers (who are fraction $(1 - \beta)$), and “naïve” consumers (who are fraction $\beta \gamma$).

Second, we introduce a period zero during which consumers are behind a “veil of ignorance”. In this period, consumers believe that they will be naïve with probability $\beta \hat{\gamma}$, high risk with probability $\beta (1 - \hat{\gamma})$, and low risk with probability $1 - \beta$. In period one, their subjective type is revealed and they fully update, completely replacing their period-zero beliefs. In period zero, a representative citizen decides whether to delegate the power to regulate the market to an independent regulator, who we model as a social planner with population weights on the objective welfare of each ex-post risk type.\textsuperscript{10} If the citizen delegates, the regulator will implement the regulator’s preferred regulatory mechanism instead of progressing through stages 1-3 (we refer to the “regulation subgame”). Otherwise, the consumers and firms interact in the free market (the “market subgame”).

3.1 Market Subgame with Three Types
Let $\overline{\alpha}_t \equiv \frac{\beta \gamma \alpha_h + (1 - \beta) \alpha_l}{1 - \beta + \beta \gamma}$ represent the average add-on risk among those believing themselves to be low-risk. The following proposition characterizes the competitive equilibria in the market subgame.

**Proposition 4.** In all subgame perfect Nash equilibria of the market sub-game,

1. If the gains from trade are low enough ($G < -k(\overline{\alpha}_t - \alpha_l)$), no contracts are accepted.

\textsuperscript{10}If the regulator is biased in favor of the naïve, the results will be identical. If he is biased in favor of the rational, the results in the $G < 0$ case would be qualitatively similar, but with different exact cutoffs. For pro-rational bias and $G > 0$, regulation has no bite, since the regulator would implement outcomes identical to the competitive market.
2. Otherwise, the zero-profit contract with maximum add-on price \( (p_A' = k, p_B' = C - \pi_l k) \) is offered, and it is accepted by all naïve and low-risk consumers.

3. In addition, for positive gains from trade \( (G > 0) \), one or more zero-profit contracts with small add-on prices will be offered and accepted by all high-risk consumers. The add-on price must be low enough that the overall contract \( (p_A'^h, p_B'^h = C - \alpha_h p_a^h) \) does not attract the low-risk consumers. \( (p_A^h \leq k \frac{\pi_l - \alpha_l}{\alpha_l - \alpha_l}) \).

4. If \( G < 0 \), no high-risk consumer accepts a contract.

In equilibrium, the high risk do not really interact with those who perceive themselves to be low risk.\(^{11}\) Firms offer a bundle of contracts that allows high-risk consumers to separate from the low-risk and naïve consumers, so that high-risk consumers choose a contract with a low add-on price, and the low-risk and naïve consumers choose the highest possible add-on price. Consumption for negative \( G \) induced by naïveté persists from the benchmark model, but now only by the citizens who believe themselves to be low risk.

The objective and subjective unregulated expected payoff of the high-risk is \( U^h = \max\{0, G\} \). For the low risk, it is \( U^l = \max\{0, G + k(\overline{\alpha}_l - \alpha_l)\} \), and the objective expected payoff to the naïve is \( U^n = (G + k(\overline{\alpha}_l - \alpha_h)) \) if \( G > -k(\overline{\alpha}_l - \alpha_l) \) (and 0 otherwise) but they believe themselves to receive the low-risk payoff.

Again, a monopoly is quite similar, with contracts that perfectly separate the types and reap all the subjective consumer surplus to the monopolist.\(^{12}\) Consumption decisions are identical to the competitive market, but no type gets any subjective consumer surplus.

### 3.2 Regulation Subgame with Three Types

In the setting with three types, a budget-balanced regulatory mechanism consists of a set of messages from the citizens \( m = (i, j) \in \{h, l\} \times \{0, 1\} \), message-contingent transfers \( (t_j, t_j^A) \in \mathbb{R}^4 \), where \( t_j \) represents the transfer paid by a consumer of type \( j \), and \( t_j^A \) is the additional transfer paid by a consumer of type \( j \) who needs the add-on, and probabilities of receiving the base good and add-on \( (\pi_j, \pi_j^A) \in [0, 1]^4 \). By

\(^{11}\)The separation of subjective risk types is because \((p_B, p_A)\) pairs are perfectly observable and contractible, unlike in some shrouded attributes models.

\(^{12}\)Formally, a monopolist would offer two contracts, \((p_B'^h = B, p_A'^h = 0)\) for the high risk and \((p_B'^l = B - k\alpha_l, p_A'^l = k)\) for the low-risk and naïve, if \( G \geq 0 \) and only the low-risk/ naïve contract if \( 0 > G \geq -k(\overline{\alpha}_l - \alpha_l) \).
the revelation principle it suffices to consider truthful, direct mechanisms. Budget balance requires that

$$\beta(1 - \gamma)[t_h + \pi_h(\pi_h^A \alpha_t t_h^A - C)] + (1 - \beta + \gamma\beta)[t_l + \pi_l(\pi_l^A \alpha_t t_l^A - C)] = 0$$  \hspace{1cm} (6)$$

Ex-post incentive compatibility requires that for each consumer type, $j$, 

$$0 \leq t_j^A \leq k.$$  \hspace{1cm} (7)$$

The ex-ante incentive compatibility constraint for the low risk and naïve is 

$$\pi_l[B - \alpha_l(\pi_l^A t_l^A + k(1 - \pi_l^A))] - t_l \geq \pi_h[B - \alpha_h(\pi_h^A t_h^A + k(1 - \pi_h^A))] - t_h,$$  \hspace{1cm} (8)$$

and that for the high risk is 

$$\pi_h[B - \alpha_h(\pi_h^A t_h^A + k(1 - \pi_h^A))] - t_h \geq \pi_l[B - \alpha_l(\pi_l^A t_l^A + k(1 - \pi_l^A))] - t_l.$$  \hspace{1cm} (9)$$

Finally, the period-1 participation constraints require that, for each type, 

$$\pi_j[B - \alpha_j(\pi_j^A t_j^A + k(1 - \pi_j^A))] - t_j \geq 0.$$  \hspace{1cm} (10)$$

We assume that the regulator has to only induce market participation in period 1, a weaker condition than the pseudo-participation constraint (3) in the benchmark model, while political support is garnered in period zero. Once delegation occurs the regulator has greater latitude for regulation, while the period zero decision effectively replaces the pseudo-participation constraint.

Assume a paternalistic regulator maximizes the weighted sum of the objective consumer surplus of the three consumer types, with a weight of $\psi_n$ on the naïve, a weight $\psi_l$ on the low-risk, and a weight $\psi_h$ on the high-risk. He solves

$$\max_{t, \pi} \left\{ \psi_n \left[ \pi_l[B - \alpha_l(\pi_l^A t_l^A + k(1 - \pi_l^A))] - t_l \right] + \psi_l \left[ \pi_l[B - \alpha_l(\pi_l^A t_l^A + k(1 - \pi_l^A))] - t_l \right] + \psi_h \left[ \pi_h[B - \alpha_h(\pi_h^A t_h^A + k(1 - \pi_h^A))] - t_h \right] \right\}$$  \hspace{1cm} (11)$$

subject to constraints (6-10).
Proposition 5. When a paternalistic regulator places weight equal to population share on each type’s welfare \(\psi_h = \beta(1 - \gamma)\) and \(\psi_n = \gamma\beta\), any optimal regulatory mechanism dictates consumption if and only if it is efficient (i.e., sets \(\pi_h = \pi_l = 1\) whenever \(G > 0\) and \(\pi_h = \pi_l = 0\) whenever \(G < 0\)).

A multiplicity of regulatory mechanisms implement the efficient outcome but with different distributional outcomes. Since the regulator is indifferent among citizens, however, he is indifferent among those mechanisms. We assume that the regulator will implement the simplest of these, where \(t^A_l = t^A_h = 0\) and \(t_l = t_h = C\) if \(\pi = 1\). This simple mechanism is always incentive-compatible and in the set of optimal mechanisms. It guarantees the same objective and subjective expected utility of all consumer types \(U^{reg} = \max\{0, G\}\). None of the qualitative results, below, change if we assume that the regulator implements some alternative optimal regulation, although the exact cutoffs may change.\(^{13}\)

3.3 Delegation Choice Behind The Veil

A period-zero citizen believes that he will be naïve with probability \(\hat{\gamma}\beta\), rationally low risk with probability \((1 - \beta)\), and rationally high risk with probability \(\beta(1 - \hat{\gamma})\). Given the payoffs derived from Proposition 4, his expected utility from the unregulated market from behind the veil is

\[
EU^{\text{unreg}} = \begin{cases} 
G + k[(1 - \beta)(\alpha_l - \alpha_l) + \hat{\gamma}\beta(\alpha_l - \alpha_h)] & \text{if } 0 \leq G \\
[\hat{\gamma}\beta + (1 - \beta)]G + k[(1 - \beta)(\alpha_l - \alpha_l) + \hat{\gamma}\beta(\alpha_l - \alpha_h)] & \text{if } -k(\alpha_l - \alpha_l) < G < 0 \\
0 & \text{if } G < -k(\alpha_l - \alpha_l).
\end{cases}
\]

His expected utility from the regulated market is \(EU^{\text{reg}} = \max\{G, 0\}\).

A citizen’s period 0 belief about the probability he will need the add-on in period 3 of the market subgame, conditional on believing himself to be low-risk ex-post is

\(^{13}\)The payoff for negative gains from trade is uniquely zero in all efficient regulations, but for positive \(G\) the regulator is indifferent among a continuum of redistributionary regulations. If the regulator chooses a regulation that involves some equilibrium transfers from the naïve to the low risk (\(0 < t^A_l < k\)), the regulated expected payoff will fall between \(G\) and the unregulated payoffs. But a behind-the-veil consumer who prefers the market payoff to \(G\) will also prefer a convex combination of \(G\) and the market payoff to \(G\), and vice-versa. He simply will not feel as strongly about his preference. Only in the extreme case \(t^A_l = k\), which exactly replicates the market payoff, will the posited citizen be indifferent.
given by
\[ \hat{\alpha}_t = \frac{(1 - \beta)\alpha_t + \beta\hat{\gamma}\alpha_h}{1 - \beta + \beta\hat{\gamma}}. \]

We can use this quantity to characterize the optimal choice of whether to delegate to a regulator.

**Proposition 6.** In every subgame perfect Nash equilibrium of the delegation game behind the veil of ignorance

1. A representative citizen with weakly pessimistic beliefs about his naiveté (\(\hat{\gamma} \geq \gamma\)) prefers to delegate for all gains from trade, with strict preference when the pessimism is strict or regulation binds (i.e., \(-k(\bar{\alpha}_t - \alpha_t) < G < -k(\bar{\alpha}_t - \hat{\alpha}_t)\) or \(\hat{\gamma} > \gamma\)).

2. A representative citizen with optimistic beliefs about his naiveté (\(\hat{\gamma} < \gamma\)) prefers to delegate if and only if the gains from trade are sufficiently small (\(G \leq -k(\bar{\alpha}_t - \hat{\alpha}_t)\)), with strict preference when regulation binds (i.e., \(-k(\bar{\alpha}_t - \alpha_t) < G < -k(\bar{\alpha}_t - \hat{\alpha}_t)\)).

In the market, low-risk and naıve consumers are willing to buy a good with negative gains from trade if they believe \(\alpha_t\) is sufficiently below the average risk for their contract choice \(\bar{\alpha}_t\). Behind the veil, this decision turns on \(\hat{\alpha} \neq \alpha_t\) because they (imperfectly) take into account the possibility that later they will be naıve. In particular, if \(\hat{\gamma} < \gamma\), consumers are optimistic about being naıve, so they delegate in cases of significantly negative \(G\), only. The range of \(G\) for which delegation occurs grows larger in sophistication (\(\hat{\gamma}\)), and smaller in the fraction naıve (\(\gamma\)), the risk differential (\(\alpha_h - \alpha_t\)), and the maximum add-on price (\(k\)). If \(\hat{\gamma} = \gamma\), consumers are sophisticated in period zero and choose regulation if and only if \(G < 0\). If \(\hat{\gamma} > \gamma\), consumers are actually pessimistic and would like to regulate even in the \(G > 0\) markets in order to avoid the cross-subsidization entirely.

The veil of ignorance represents the idea that consumers are sophisticated about the possibility that they could be biased in some dimensions but are uncertain of which biases they have exactly.\(^{14}\) Consumers have second-order sophistication – they are naıve about their bias in each particular dimensions but sophisticated insofar as they recognize that they be biased along some dimensions. To the extent that citizens are more correct in their beliefs about their riskiness as they are more “distant” from

\(^{14}\)We thank an anonymous referee for this suggestion.
the market transaction, they will support better regulatory decisions. They may, for example, be willing to establish an agency staffed by behavioral experts such as the CFPB or the UK Behavioral Insights Team, an approach that has proven somewhat popular.\textsuperscript{15} Our distinction between general and specific contexts matches a common finding in the literature on overconfidence, that overconfidence is smaller in estimates across a set of items compared to estimates of a single item (Sniezak and Buckley, 1991; Moore and Healy, 2008). Regulatory delegation serves as a commitment device, but not the usual kind in which sophisticated consumers commit to a future consumption profile. Instead, citizens give up general political power because they believe they will misuse it in specific instances in which they will be biased and naïve.

Small amounts of altruism can have an impact behind the veil. A partially-altruistic citizen who puts some weight $\delta$ on the welfare of other citizens from behind the veil, relative to a weight of 1 on his own welfare, will consider regulation more favorably. This altruist can be modeled as either a utilitarian, who values the objective welfare of the average citizen, or as a Rawlsian, who values the welfare of the worst off member of society (here, the naïve). Both formalizations of altruism move upward the cutoff below which an imperfectly sophisticated citizen prefers delegation (part 2 of Proposition 6), although the Rawlsian altruist moves more. Rawlsian altruism can also induce support for regulation that shuts down redistribution: specifically, when the concern for the worst off is strong enough ($\beta(\gamma - \hat{\gamma}) \leq \delta$), a Rawlsian altruist supports regulation for all $G$, including $G > 0$.

Finally, in our simple model, the consumer behind the veil would always weakly prefer delegating to the regulator to allowing a monopolist to operate in the free market. The best any type does under the monopoly is zero consumer surplus, and the naïve do even worse. If we introduced a third choice, however, of a regulator whose sole job was to induce competition, then that choice would also always dominate a free-market monopolist and the choice between the pro-competition regulator and the paternalistic regulator would exactly mirror the choice outlined in Proposition 6, with the paternalistic regulator preferred when delegation would be preferred.

\textsuperscript{15}“The Nudgy State: How five governments are using behavioral economics to encourage citizens to do the right thing” Joshua E. Keating Foreign Policy, Jan. 2, 2013
4 Regulation and The Credit Card Industry

As an illustration, consider the market for credit cards. The base good of credit cards is transactional – the capacity to make purchases without carrying cash. The base-good price is small or negative: any annual fee minus the value of any rewards.

Credit cards have (at least) two varieties of optional services that can be described as add-ons in our model, “gotchas” and “indebtedness”. Each of these add-ons has a vector of characteristics \( (G_j, k_j, \alpha_{l,j}, \alpha_{b,j}, \beta_j, \gamma_j) \), indexed by \( j \in \{g, i\} \). Many consumer seem naïve about their ability to avoid these these add-ons.\(^{16}\)

The first type of add-on, which we refer to as “gotchas”, includes significant fees and interest rate increases triggered by contingencies: over-the-limit and late-payment fees, universal default, double-cycle billing. Contracts could be offered that have more or fewer restrictions on card usage and balance payments, and marginal net gains from trade from a contract with many “gotchas” is probably negative \( (G_g < 0) \), yet it will be accepted by the naïve and low-risk consumers in a competitive equilibrium if the cross-subsidy available is sufficiently large \( (G_g > k(\bar{\alpha}_{l,g} - \alpha_{l,g})) \).\(^{17}\)

A second type of add-on, which we refer to as “indebtedness”, includes the ability of card holders to make purchases that they do not have enough liquid funds to afford. The marginal net gains-from-trade from this add-on is probably larger \( (G_i > G_g) \). If it is actually positive, it would be accepted by all consumer types, although the different subjective risk types would choose a different contracts.\(^{18}\)

In analyzing the parameter values of these two kinds of add-ons, we rely on Stango and Zinman (2009), who provide a rich summary of every financial transaction of 917 households over a two year period that suggests the approximate magnitudes of the risk characteristics of our two add-ons.\(^{19}\) In these data, the add-on price is larger for

\(^{16}\)Shui and Ausubel (2005) find that most consumers, when offered either an introductory offer with a very low rate and short duration, or an offer with a higher rate and longer duration, took the shorter length offer, despite in most cases losing money because of it. Similarly, Agarwal et al. (2006) find that when choosing between an offer with a fee and a lower rate or no fee and a higher rate, many borrowers took the offer with no fee despite the fee offer being less costly ex post.

\(^{17}\)The fundamental behavioral problem for gotchas is probably unawareness (as in Gabaix and Laibson (2006)), as naïve consumers will underestimate how often their behavior will trigger such a contingency.

\(^{18}\)The fundamental behavioral problem for indebtedness is present-biased preferences (as in DellaVigna and Malmendier (2004) or Heidhues and Koszegi (2010)), as naïve consumers will overestimate their future selves’ willingness to pay down balances. Meier and Sprenger (2010) find that card-holders who make present-biased choices in incentivized choice experiments have significantly higher outstanding balances.

\(^{19}\)These households are slightly more educated and have slightly higher incomes than the national average, but for comparing add-on characteristics, it seems likely that the bias from this selection
indebtedness than for gotchas (i.e., \(k_i > k_g\)). Among consumers who pay fees, the 25th percentile fee monthly cost is $3.88 and the 75th percentile is $29.50. For interest charges, the 25th percentile monthly cost is $6.73 while the 75th percentile is $76.16. Next, the risk differential between the high-risk and low-risk is probably larger for indebtedness than for gotchas (\(\alpha_{h,i} - \alpha_{l,i} > \alpha_{h,g} - \alpha_{l,g}\)), as measured by the month to month correlation in costs for a given household. The median month-to-month correlation in fees paid is -0.06. For interest payments, the median correlation is 0.46. Taken together, these two characteristics imply that the cross-subsidization from gotchas is smaller than that from indebtedness. We have no strong prior or evidence about which sort of add-on induces the higher levels of naïveté, but for similar levels of naïveté and sophistication over that naïveté, the representative citizen would accept lower marginal gains from trade from indebtedness before delegating to the regulator than he would for the gotchas.

**Regulation under The CARD Act** The Credit CARD Act of 2009 placed restrictions on what credit card companies can charge for some of these add-ons. It capped over-the-limit and late-payment fees. Interest rates can not increase during the first year of having the card. Forty-five day advanced notification for APR changes is mandated. Universal default, the practice of increasing interest rates based on payment histories from unrelated accounts, such as utilities or other credit cards, is banned. These changes essentially limit surprises about fees and interest charges, corresponding to limits on gotcha add-ons (see Bar-Gill and Bubb (2012) for a summary of the CARD Act’s provisions).

Consistent with the CARD Act having regulated mainly gotchas, Bar-Gill and Bubb (2012) and Agarwal et al. (2013) document that since the Act took effect, prohibited fees have declined. Profits of credit card issuers have fallen, and Agarwal et al. find that add-on fees which were previously kept by firms as profit are now retained by consumers, so the CARD Act helped at least some consumers.

**Politics of The Act, Its Timing, and Form** While the Act passed Congress with widespread bipartisan support, seven members of the House Financial Services Committee dissented because “in limiting credit card issuers’ ability to price for risk, Congress needs to avoid overreacting by forcing responsible card-holders to subsidize irresponsible ones through higher fees and fewer rewards” (in our model, an increase issue is small.
Consumer anxiety about prices was evident in news articles about the CARD Act, which routinely observed that fees could increase, rewards fall, or credit tighten. These anxieties are sometimes well-founded. In a slightly different consumer financial market, Stango and Zinman (2011) find that stricter Truth in Lending Act enforcement raised interest rates for sophisticated consumers without reducing them for naïve consumers, because when required to make contracts transparent, finance companies eliminated contractual terms that now would clearly dominate their more profitable contracts.

Despite these concerns, the Act was passed. The analysis in Section 3 suggests some plausible reasons for the timing of its passage. The recession of December 2007 to June 2009 may have made the risks of these add-ons more salient for credit card users, raising \( \hat{\gamma} \) for both add-ons. The recession also exposed otherwise low-risk individuals to more financial risk, possibly reducing the risk differential \( (\alpha_h - \alpha_l) \). Either of these parameter shifts could trigger regulation in the behind-the-veil model, as they both make cross-subsidization less attractive.

Our model predicts that the CARD Act would take the form of restrictions on gotchas rather than restrictions on indebtedness. If \( G_i > 0 \), reform of indebtedness is unappealing unless consumers are paranoid \( (\hat{\gamma} > \gamma) \). In addition, if \( k_g < k_i \) and \( (\alpha_h - \alpha_l)_g < (\alpha_h - \alpha_l)_i \), the perceived value of cross-subsidization from gotchas is relatively low, leaving less room when \( G_g < 0 \) for the cross-subsidization to outweigh the inherent negative value of gotchas. While there was considerable public support for the Act, that support was limited to fees and interest rate surprises, i.e., gotchas.

Some commentators have called for additional regulations that target indebtedness. These proposals would limit issuers’ ability to directly target consumers who overestimate their ability to repay the revolving credit. For example, Bar-Gill and Bubb (2012) propose a floor on interest rates, and Slowik (2012) proposes increased minimum monthly payments. We believe these proposals will be less politically feasible. Indeed, during consideration of the Act, Senator Bernie Sanders tried to amend

---

22 An additional reason that consumers might oppose regulation of interest rates (suggested to us by an anonymous referee) is that high interest rates can discipline their borrowing, potentially serving as a commitment device.
the CARD Act to include an interest rate cap, but his proposal was rejected by the Senate 33-60 (S.A. 1062 to S.A. 1058 to H.R. 627).

5 Conclusions and Implications

Our model illustrates an important feature of markets with naïve consumers – a disinclination by participants towards regulation that would eliminate avoidable dead-weight losses. The disinterest in behavioral regulation may actually exceed that of more traditional economic regulation. When regulating an externality or monopolist, for example, there are winners and losers, but the total expected gains of efficient regulation to the winners is bigger than the expected loss to the losers, so in an idealized polity in which everyone expresses their self-interest politically, traditional economic regulation would be efficient. Regulation to address inefficiencies induced by naïve consumers is different because the losers misperceive their self-interest, leading them to oppose efficient regulation.

We demonstrate these results under specific assumptions about the underlying market and type of consumer bias, but we believe the intuition applies more broadly to many models that have been used to derive behavioral policy prescriptions. DellaVigna and Malmendier (2004), for example, consider two-part contracts for consumption decisions with costs and benefits separated in time, such as going to the gym or using a credit card. With partially naïve quasi-hyperbolic discounters, there is scope for welfare-improving regulation (their Proposition 3). In perfectly competitive markets, though, all consumers would oppose the efficient regulation, since the equilibrium prices maximize their (subjective) consumer surplus. Similarly, in Gabaix and Laibson’s (2006) shrouded-attribute model, regulations that inform the myopic consumers of the existence of the add-on would benefit them, but such an intervention would hurt the sophisticated consumers, who benefit from the subsidized base good. From a purely self-interested perspective, they would oppose it. The policy preferences of the myopic consumers are not well-defined, but they are unlikely to strongly demand regulation for a good that they do not know exists. Finally, consider Heidhues and Koszegi (2010), in which naïve quasi-hyperbolic discounters agree to contracts with

24 We do not want to overstate the efficiency of the political process. It is fraught with rent-seeking. Not all groups are equally organized, and there many examples of inefficient traditional economic regulation. Our point is that, abstracting away from all purely political problems, the pressure should be there to do the efficient thing. After all, this is the very definition of dead-weight loss.
large penalties. Limiting large penalties for deferring small payments in credit markets would be welfare improving (on net), because it allows naïve consumers to better anticipate their ex-post repayment decisions. When firms are not able to discriminate between sophisticated and naïve consumers, they leave some rents to the sophisticates in order to take advantage of the naïve (their Proposition 7). Ex ante, both consumer types believe that the regulation makes them worse off, although the naïve are wrong in that belief.

Because behavioral regulation involving naïve consumers is unlike traditional economic regulation, some force beyond consumer self-interest is necessary for first-best policies to be robustly enacted. We can think of at least two responses to this fact.

First, one could imagine some welfare-maximizing regulator who is insulated from political pressures. We take one step in this direction in our analysis of the support for regulation “behind the veil.” But fully insulated regulators are risky, and this response leaves the citizens with a commitment problem, since the citizens would like to undo the regulation ex-post. An opportunistic politician might take advantage of this desire to change the regulation ex-post. This political unraveling of commitment devices is the political failure that Bisin et al. (2013) explore.

Second, a small amount of altruism might suffice to induce some support for efficient regulation, especially if consumers are partially sophisticated. In the behind-the-veil model, regulatory authority is delegated, there is a mass of high-risk individuals who are indifferent about the way surplus is divvied up between the low risk and the naïve. If they care somewhat about the payoffs of the other members of society, however, they would have a strict preference for efficient regulation. If the high risk are pivotal, their altruism could help solve some of the commitment problems of sticking to the efficient regulation.
6 References


Spiegler, Ran, Bounded Rationality and Industrial Organization, Oxford University Press, 2011.


A Proofs

A.1 Proof of Proposition 1

Assuming firms can only offer a countable number of contracts, a strategy profile for firms in this game consists of a vector of contracts \( P = (p_B, p_A, \rho)_{ij} \), where firm \( i \) may offer several contracts (indexed by \( j \)) and \( \rho_{ij} \) is the probability that \( i \) offers its \( j \)th contract. A strategy profile by consumers is a vector of contract-vector dependent consumers’ base-good consumption choices, \( C^B(P) = C^B(i, j)_m(P) \), together with an contract and add-on need dependent add-on consumption choice vector \( C^A(C^B, P, N) \), where \( C^B(i, j)_m \geq 0 \) represents the probability that consumer \( m \) chooses contract \( j \) from firm \( i \), and \( C^B(0, 0)_m \geq 0 \) represents the probability that consumer \( m \) chooses not to consume. Of course, we require that \( \sum_{i,j} C^B(i, j)_m = 1 \).

A sub-game perfect equilibrium requires the consumers to be expected welfare maximizing at each history and the firms to be profit-maximizing at the initial node of the game (the only time they act). Indicate such a sub-game prefect equilibrium by \( (P^*, C^B^*(P), C^A^*(P, C^B, N)) \).

Beginning with the add-on consumption decision, welfare-maximization requires that \( C^A^*(N, P = 1) = 1 \) whenever the add-on price for the selected contract is below \( k \) and \( C^A^*(N, P = 1) = 0 \) whenever it is above \( k \). It also requires that \( C^A^*(N, P = 0) = 0 \) unless \( p_A \leq 0 \).

Next move to the second-period subgames, where consumers choose among the contracts. The consumers must choose whichever contract gives them the lowest expected price, as long as that contract has expected price below \( B \). If indifferent among a set, any distribution among that set is consistent with subgame perfection.

A consumer who believes he is type \( \alpha \) has expected price from contract \( (p_B, p_A) \) of \( e[p] = p_B + \alpha p_A \) if \( p_A \leq k \) and \( e[p] = p_B + \alpha k \) if \( p_A \geq k \).

Moving back to the subgames of the first period of the game, contract choice by the firms, we first establish some conditions that every equilibrium vector of contracts must satisfy.

Lemma 1. Any contract accepted in equilibrium with positive probability must deliver zero profits.

Proof. Let \( (p_B, p_A) \) be a contract chosen by some consumer with positive probability. It must yield non-negative expected profits, or the offering firm would quit offering it.
and increase profits. It must yield non-positive profits, or else any firm not currently
selling to any customers could offer \((p_B - \epsilon, p_A, 1)\), attract some consumers for sure, and
make positive profits. If all firms have contracts accepted with positive probability,
then the firm making positive profits could cut its own base-good price by \(\epsilon\), holding
its add-on price fixed, attract more consumers, and increase profits. ■

**Lemma 2.** If \(\gamma > 0\), any contract accepted in equilibrium with positive probability
must have an add-on price of \(k\).

*Proof.* Let \((p_B, p_A)\) be a contract chosen with positive probability by some consumer
in equilibrium. We know from the forgoing Lemma that it delivers zero expected
profits. This contract must have \(p_A = k\), since any firm could deviate to \((p_B - \alpha_i \delta -
\epsilon, p_A + \delta)\) attract all consumers and make positive profits, for small enough \(\delta\) and
\(\epsilon\). The deviator collects \(\alpha \delta\) more in add-on fees and \(\alpha_i \delta + \epsilon\) less on base good sales,
while the consumer expects to pay \(\epsilon\) less, in expectation. Since some consumer was
choosing the original contract with positive probability, this new contract will now
attract all consumers. ■

**Lemma 3.** If a consumer of risk type \(\alpha\) accepts any contracts in equilibrium with
positive probability, then there must be contracts offered with probability 1 by at least
two different firms such that the consumer is indifferent between the two contracts
and the one(s) he is accepting.

*Proof.* Because consumers of perceived risk type \(\alpha\) accept contracts in equilibrium,
they must only accept zero-profit contracts, so one or more firms are offering zero-
profit contracts with probability greater than zero. If no firms offer zero-profit con-
tacts to that type with probability one, then any firm \(i\) offering a zero-profit contract
\(j\) to that type can increase their expected profits by raising \((p_B)_{ij}\), as there is some
probability no other contracts will be offered. Likewise, if only one firm \(i\) offers a
zero-profit contract \(j\) to that type with \(\rho_{ij} = 1\), then firm \(i\) can increase its expected
profits by raising \((p_B)_{ij}\). ■

These lemmas suffice to characterize equilibrium contract offers and acceptances,
but not to pin down exactly which firms offer which contracts or which consumer
buys from which firm.

Consider the case when \(G > k(\bar{\alpha} - \alpha_i)\). Taking the first two lemmas together,
any equilibrium with a contract accepted must have all accepted contracts of the
form outlined in the proposition \((p_B^* = C - \bar{\alpha}k, p_A^* = k)\). This contract yields a
consumer believing himself to be type \( \alpha_l \) an expected total price of \( C + k(\alpha_l - \overline{\alpha}) \), and this price is sufficient to induce consumption whenever \( G > k(\alpha_l - \overline{\alpha}) \). To show that such an equilibrium exists, consider the case where every firm offers the \((p_B^* = C - \overline{\alpha}k, p_A^* = k)\) contract, making zero expected profits whether they sell a base good or not. A deviation by a firm to \((p_B', p_A')\) that makes strictly positive profits must have \( p_B' > C - \overline{\alpha}p_A' \). This deviation yields an expected price to the consumer of 
\[
e[p'] = p_B' + \alpha_l p_A' > C - \overline{\alpha}p_A' + \alpha_l p_A' = C + p_A'(\alpha_l - \overline{\alpha}).
\]
This expected price decreases in \( p_A' \), so it is always above the equilibrium expected price, and thus there is no strictly profitable deviation. Furthermore, an equilibrium with no accepted contracts is not possible, since a deviation to a contract of the type \((p_B = C - \overline{\alpha}k + \epsilon, p_A = k)\) would be both acceptable to all consumers and profitable, for sufficiently small \( \epsilon \).

Now consider the case when \( G < k(\alpha_l - \overline{\alpha}) \). By the analysis above there can not be an equilibrium in which a contract is accepted. Furthermore, it is an equilibrium for no contract to be accepted. Consider any set of contracts, none of which are attractive enough to induce consumers to buy. Any profitable deviation must induce some consumer to buy, so must have expected price below consumption utility \( B \). Given the analysis in Lemma 2, the most profitable contract with expected price below \( B \) is \((p_B = B - \alpha_l k - \epsilon, p_A = k)\). This contract would attract all consumers and yield the firm expected profit of \( G + k(\overline{\alpha} - \alpha_l) - \epsilon < 0 \).

### A.2 Proofs of Proposition 2 and Proposition 3

Substituting for the budget-balance condition we can rewrite the problem in (4) as
\[
\max_{t, \pi} \left\{ G + t^A[\overline{\alpha} - \psi_n \alpha_n - \psi_l \alpha_l] + (\pi^A - 1)k[\psi_h \alpha_h + \psi_l \alpha_l] \right\}. \tag{13}
\]
Ignore, for now, the consumers’ participation constraints. Essentially, there are two regulatory approaches: \( \pi = 0 \) and \( \pi = 1 \), and the regulator will choose the better of the two. In the \( \pi = 0 \) case, \( \pi^A \) and \( t^A \) are moot, budget-balance requires \( t = 0 \), and the regulator achieves a payoff of zero. If \( \pi = 1 \), \( \pi^A \) weakly increases the regulator’s payoff and only loosens the constraints, so it will be set to 1. \( [\overline{\alpha} - \psi_n \alpha_n - \psi_l \alpha_l] \gtrless 0 \) whenever \( \psi_n \lessgtr \gamma \). Thus, when \( \psi_n < \gamma \), the regulator’s payoff increases in \( t^A \), so it will be set to the largest value consistent with (2). This approach yields a payoff that exceeds the \( \pi = 0 \) case whenever \( G > -k[\overline{\alpha} - \psi_n \alpha_n - \psi_l \alpha_l] \). When \( \psi_n = \gamma \), the regulator’s payoff is independent of \( t^A \), so the regulator will prefer this approach for \( G > 0 \). When \( \psi_n > \gamma \), the regulator’s payoff decreases in \( t^A \). Since (2) requires
If \( t^A \geq 0 \), the best the regulator can do is set \( t^A = 0 \). So, again, the regulator prefers this approach to \( \pi = 0 \) whenever \( G > 0 \).

Return, finally, to the participation constraints. If we require only (5), the constraint never binds. Obviously, when \( \pi = 0 \) and \( t = 0 \), the consumer makes expected payoff of exactly zero. Whenever \( \pi > 0 \) and \( t^A = 0 \), the budget-balance condition assures that consumers earn consumer surplus of \( G \), which is non-negative for the parameters leading to \( \pi > 0 \). Finally, whenever \( \pi > 0 \) and \( t^A = k \), the budget-balance condition guarantees that consumers earn subjective expected consumer surplus of \( G + k[\bar{\pi} - \alpha_l] \) which is non-negative for the parameters leading to \( \pi > 0 \).

For the stricter participation condition (3), substituting budget balance yields the following condition

\[
\pi[G + (\bar{\alpha} - \alpha_l)t^A] \geq \max\{0, G + (\bar{\alpha} - \alpha_l)k\}. \tag{14}
\]

For \( G + (\bar{\alpha} - \alpha_l)k > 0 \), this can be satisfied only by \( \pi^A = 1, t^A = k, \pi = 1 \), exactly the free market outcome. \( G + (\bar{\alpha} - \alpha_l)k < 0 \), it can be satisfied only with \( \pi = 0 \), exactly the free-market outcome.

### A.3 Proof of Proposition 4

We divide the vector of contract choices into three types, depending on how the consumers divide amongst the contracts. First, there may be a separating equilibrium, in which no two consumers who believe themselves to be different types choose the same contract. Second, there may be a pooling equilibrium, in which all consumer types choose the same contract. Finally, there may be a semi-separating equilibrium, in which there is a contract that is chosen by only one type and a contract which is chosen by multiple types.

The basic structure of this proof is similar to that of Proposition 1. The notational formalities exactly mirror those from that proof, and Lemmas 1 (any contract accepted in equilibrium is zero-profit) and 3 (at least two firms offer zero-profit contracts with probability one to any perceived risk type that accepts any contracts) continue to hold. We now establish some additional conditions that every equilibrium vector of contracts must satisfy in the three-type case and finally we characterize the parameter combinations for which each type of equilibrium exists.

**Lemma 4.** Any contract accepted in equilibrium with positive probability by some consumers of all types must have an add-on price of 0.
Proof. Let \((p_B, p_A)\) be a contract chosen by some consumers who believe themselves to be low risk and some high-risk consumers. It must have \(p_A = 0\) or else an entrant could offer \((p_B + \delta \alpha_h - \epsilon, p_A - \delta)\), attract all the high risk, and make greater profits than the original contract yields.

**Lemma 5.** If \(\gamma > 0\), any contract accepted in equilibrium with positive probability only by consumers who believe themselves to be low risk must have an add-on price of \(k\).

**Proof.** Let \((p'_B, p'_A)\) be a contract chosen by some consumers believing themselves to be low-risk. This contract must have \(p'_A = k\), since otherwise a deviation to \((p_B - \alpha_l \delta - \epsilon, p'_A + \delta)\) for small enough \(\delta\) and \(\epsilon\), attracts (at least) all those believing themselves to be low risk and yield positive profits. The deviator collects \((\bar{\alpha}_l \delta)\) more in add-on fees and \(\alpha_l \delta + \epsilon\) less on base good sales.

Now we characterize the subgame perfect equilibria of the game, focusing on each type in turn.

**Separating:** In a separating equilibrium, there are (at least) two contracts \((p'_h, p'_h)\) and \((p'_B, p'_A)\) chosen by high-risk and low-risk/naive consumers, respectively.

From Lemma 1, \((p'_B, p'_A)\) yields zero profits, so

\[
p'_B + \bar{\alpha}_l [p'_A] = C.
\]

Given this base good price, the expected utility of those believing themselves to be low risk taking this contract is

\[
EU_l = G + (\bar{\alpha}_l - \alpha_l)p'_A.
\]

If there are no naïve consumers, \(\bar{\alpha}_l = \alpha_l\) and this is simply \(G\). If there are any naïve consumers, this contract must have \(p'_A = k\), from Lemma 5. So when \(\gamma > 0\), \(p'_A = k\) and \(p'_B = C - \bar{\alpha}_l k\) and \(EU_l = G + (\bar{\alpha}_l - \alpha_l)k\). The low risk/naive are willing to buy this contract as long as \(G > -(\bar{\alpha}_l - \alpha_l)k\).

Now consider the high-risk contract, from Lemma 4, \(P^h_B = C - \alpha_h(P^h_A)\), and \(EU_h = G\). For it to not attract the low risk, the expected price to the low risk under it must be higher than under the low-risk contract, i.e., \(p^h_B + \alpha_l(p^h_A) \geq C - (\bar{\alpha}_l - \alpha_l)k\).
Satisfying these conditions together requires

$$(\overline{\alpha}_l - \alpha_l)k \geq (\alpha_h - \alpha_l)p_A^h.$$ 

Multiple $p_A^h$’s may satisfy this condition, but $p_A^h = 0$ always will. The high risk are willing to buy their contract as long as $G > 0$. If they do not take this contract, there is no separating contract they take.

Finally, think about a separating equilibrium in which the high risk do not consume. If they take the low-risk/naive contract, they receive $EU_h = G + k(\overline{\alpha}_l - \alpha_h)$. If this is not sufficient to attract the high risk, an entrant can offer $p_B^h = B - \epsilon$ and $p_A^h = 0$ and attract the high risk. Regardless of whether it also attracts the low-risk/naive, this contract makes a profit whenever $G > 0$.

A separating equilibrium where the high-risk consume and low-risk/naive do not can only exist in the knife-edged case where $G = 0$ and $\gamma = 0$, since any contract is at least as attractive to the low-risk as it is to the high-risk, and a contract setting $p_A = k$ and $p_B = C - \overline{\alpha}_l k$ is even more attractive than that, whenever $\gamma > 0$.

**Pooling**  In pooling equilibria, either all types take is a single contract $(p_B,p_A)$ or no types consume. Consider first pooling where all types accept the same contract. From Lemma 1, it must be zero profit. Furthermore from Lemma 4, it must have $p_A = 0$. So this contract must have $p_B = C$, which gives all types expected surplus $G$. A pooling contract does not exist if this contract produces $G < 0$. Even if the participation constraint is satisfied, an entrant at $(p_A' = k, p_B')$ attracts the low risk and naive if $p_B' + k\alpha_l < C$. Such an entrant earns profits $p_B' + \overline{\alpha}_l k - C$, so a profitable deviation exists if $(\overline{\alpha}_l - \alpha_l)k > 0$. Thus, pooling contract equilibria do not exist.

Now consider pooling where no types accept contracts. For this to be the case, firms must not be able to offer a contract that makes non-negative profits that any type accepts. Any contract is weakly more attractive to the low risk/naive, so consider an offer $p_B' = B - \alpha_l k - \epsilon$, $p_A' = k$. The profit from this contract is $B + (\overline{\alpha}_l - \alpha_l)k - \epsilon - C$. If $G < -k(\overline{\alpha}_l - \alpha_l)$, pooling on no contract is the only equilibrium.

**Semi-Separating:** Semi-separation can only happen in the trivial knife-edged case when $G = -k(\overline{\alpha}_l - \alpha_l)$ and the low-risk/naive are indifferent between consuming and not while the high-risk never consume.
A.4 Proof of Proposition 5

Solving for $t_h$ in the budget-balance condition (6),

$$t_h = \pi_h [C - \alpha_h \pi_h A^t h] + \left(\frac{1 - \gamma + \gamma^\beta}{\gamma (1 - \beta)}\right)[\pi_l (C - \alpha_l \pi_l A^t l) - t_l],$$

and substituting, we can rewrite the problem in (11) as

$$\max_{t, \pi} \left\{ \psi_h \left[ \pi_h [G - \alpha_h k (1 - \pi_h A)] - \Gamma [\pi_l (C - \alpha_l \pi_l A^t l) - t_l] \right] + \psi_l \left[ \pi_l [B - \alpha_l (\pi_l A^t l + k (1 - \pi_l A))] - t_l \right] + \psi_n \left[ \pi_l [B - \alpha_l (\pi_l A^t l + k (1 - \pi_l A))] - t_l \right] \right\},$$

where $\Gamma \equiv \frac{1 - \gamma + \gamma^\beta}{\gamma (1 - \beta)}$. Regardless of weights, the objective is independent of $t_h A^t h$, while increasing it only tightens the constraints, so we will begin by setting $t_h A^t h = 0$. Making these substitutions, we can usefully rewrite the ex-ante IC constraint the low-risk and naïve as

$$\pi_l [B - \alpha_l (\pi_l A^t l + k (1 - \pi_l A))] - t_l - \pi_h [G - \alpha_h k (1 - \pi_h A)] + \Gamma [\pi_l (C - \alpha_l \pi_l A^t l) - t_l] \geq 0,$$ \hspace{1cm} (16)

and that for the high risk becomes

$$\pi_h [G - \alpha_h k (1 - \pi_h A)] - \Gamma [\pi_l (C - \alpha_l \pi_l A^t l) - t_l] + t_l - \pi_l [B - \alpha_l (\pi_l A^t l + k (1 - \pi_l A))] \geq 0.$$ \hspace{1cm} (17)

Finally, the basic participation constraint for the high-risk becomes

$$\pi_h [G - \alpha_h k (1 - \pi_h A)] - \Gamma [\pi_l (C - \alpha_l \pi_l A^t l) - t_l] \geq 0,$$ \hspace{1cm} (18)

while that for the low-risk and naïve remains

$$\pi_l [B - \alpha_l (\pi_l A^t l + k (1 - \pi_l A))] - t_l \geq 0.$$ \hspace{1cm} (19)

and never binds, since it is implied by the low-risk IC constraint and the high-risk IR constraint. Of course, we still require the ex-post IC-conditions $0 \leq t_j^A \leq k$, and the probability constraints $0 \leq \pi_j^A \leq 1$ and $0 \leq \pi_j \leq 1$. To set up this constrained maximization, put Lagrange multipliers $\lambda^h$, $\lambda^l$, and $\lambda^IR$ on each optimizing constraint, multipliers $\tau_j^A$ and $\overline{\tau}_j^A$ on the ex-post transfers, and multipliers $\theta_j^A$, $\overline{\theta}_j^A$, $\theta_j$, $\overline{\theta}_j$ on the
probability constraints. The Kuhn-Tucker conditions are then:

\[
t_l : \Gamma \psi_h - \psi_l - \psi_h + (1 + \Gamma)(\lambda^h - \lambda^l) + \lambda^{IR} \Gamma = 0
\]

\[
t_i^A : \pi_i \pi_i^A [\psi_h \Gamma \bar{\alpha}_l - \psi_l \alpha_l - \psi_i \alpha_i - \lambda^l(\alpha_l + \Gamma \bar{\alpha}_i) + \lambda^h(\alpha_h + \Gamma \bar{\alpha}_i) + \lambda^{IR} \Gamma \bar{\alpha}_i] + \tau_i^A - \bar{\tau}_i^A = 0
\]

\[
\pi_h : (\psi_h + \lambda^h + \lambda^{IR})(G - \alpha_h k(1 - \pi_h^A)) - \lambda^l G - \alpha_i k(1 - \pi_i^A) + \theta_h - \theta_l = 0
\]

\[
\pi_i^A : \pi_i^A k(\alpha_h (\psi_h + \lambda^h + \lambda^{IR}) - \alpha_i \lambda^l) + \theta_h^A - \theta_l^A = 0
\]

\[
\pi_l : (\lambda^l - \lambda^h - \lambda^{IR} - \psi_h) \Gamma (C - \bar{\pi}_l \pi_l^A t_l^A) + (\psi_l B + \psi_n B + \lambda^l - \lambda^h)
\]

\[
- (\pi_i^A t_i^A + (1 - \pi_i^A) k) [\psi_i \alpha_i + \psi_n \alpha_n + \lambda^i \alpha_i - \lambda^h \alpha_h] + \theta_i - \theta_l = 0
\]

\[
\pi_i^A : \pi_i^A [(k - t_l^A) [\psi_i \alpha_i + \psi_n \alpha_n + \lambda^i \alpha_i - \lambda^h \alpha_h] + t_l^A \Gamma \bar{\alpha}_i (\psi_h - \lambda^l + \lambda^h + \lambda^{IR})] + \theta_i^A - \theta_l^A = 0
\]

(20)

Consider the case where \( \psi_h = \gamma (1 - \beta) \), so the regulator weighs the high-risk equal to their share of the population. Assume, for now, none of the incentive constraints bind \((\lambda^h = \lambda^l = \lambda^{IR} = 0)\). Making this substitution the \( \pi_h^A \) KT becomes \( \pi_h^A k \psi_h + \theta_h^A - \theta_l^A = 0 \), so \( \pi_h^A = 1 \) whenever \( \pi_h > 0 \). The \( \pi_h \) KT becomes \( \psi_h G + \theta_h - \theta_l = 0 \), so \( \pi_h = 1 \) whenever \( G > 0 \) and \( \pi_h = 0 \) whenever \( G < 0 \). The \( t_i^A \) KT condition becomes \( \pi_i \pi_i^A (\gamma \beta - \psi_n) (\alpha_h - \alpha_l) + \tau_i^A - \bar{\tau}_i^A = 0 \).

We assume that \( \psi_n = \gamma \beta \), so any \( t_i^A \in [0, k] \) is consistent with maximization. The \( \pi_i^A \) KT condition is \( \pi_i^A [(k - t_i^A)(\psi_i \alpha_i + \psi_n \alpha_n) + t_i^A \bar{\alpha}_i \Gamma + \theta_i^A - \bar{\theta}_i^A = 0] \), which results in \( \pi_i^A = 1 \), for any \( t_i^A \). Given this, the \( \pi_l \) KT condition is \( (1 + \gamma \beta)(G + \theta_l - \theta_i^A) = 0 \), and \( \pi_l = 1 \) whenever \( G > 0 \) and \( \pi_l = 0 \) whenever \( G < 0 \). Now, check the incentive conditions whenever \( G > 0 \). The IC-l and IC-h conditions reduce to

\[
C - t_i^A \left( \frac{\alpha_h + \Gamma \bar{\alpha}_i}{1 + \Gamma} \right) \leq t_l \leq C - t_i^A \left( \frac{\alpha_l + \Gamma \bar{\alpha}_i}{1 + \Gamma} \right),
\]

(21)

while the IR-h condition becomes

\[
t_l \geq C - \bar{\alpha}_l t_i^A - \frac{G}{\Gamma}.
\]

(22)

For \( G > 0 \), all three are satisfied at \( t_i^A = 0 \) and \( t_l = C \). As \( t_i^A \) increases, a range of \( t_l \)'s will satisfy it, and the regulator is indifferent among all of them (since they are pure transfers).
A.5 Proof of Proposition 6

When $G < 0$, the optimal mechanism adopted by a regulator would set $\pi_h = \pi_l = 0$ and consumers of every type get 0. When $G > 0$, an optimal mechanism would set $\pi_h = \pi_l = 1$, $\pi^A_l = 1$, and $t^A_l = t^A_h = 0$. In the later case, every type gets $G$.

In the unregulated market, the period-zero expected utility of entering the market sub-game outcome is

$$U_{mkt} = \begin{cases} G + k[(1 - \beta)(\bar{a}_l - a_l) + \beta\gamma(\bar{a}_l - a_h)], & \text{if } G \geq 0 \\ [1 - \beta + \beta\gamma]G + k[(1 - \beta)(\bar{a}_l - a_l) + \beta\gamma(\bar{a}_l - a_h)], & \text{if } 0 > G \geq -k(\bar{a}_l - a_l) \\ 0, & \text{if } -k(\bar{a}_l - a_l) > G. \end{cases}$$

(23)

In primitives,

$$k[(1 - \beta)(\bar{a}_l - a_l) + \beta\gamma(\bar{a}_l - a_h)] = \frac{k\beta(1 - \beta)(\bar{a}_l - a_l)}{\beta\gamma + 1 - \beta}(\gamma - \hat{\gamma}),$$

so for $G \geq 0$, $U_{mkt} \geq U_{reg}$ if and only if $\gamma \geq \hat{\gamma}$. For $-k(\bar{a}_l - a_l) > G$ the unregulated and regulated subgames both deliver expected utility of zero. In the intermediate region ($0 > G \geq -k(\bar{a}_l - a_l)$), $U_{mkt} \geq 0$ if and only if

$$G \geq -\frac{k\beta(1 - \beta)(\alpha_h - \alpha_l)}{[\beta\gamma + 1 - \beta][\beta\hat{\gamma} + 1 - \beta]}(\gamma - \hat{\gamma}) = -k(\bar{a}_l - a_l).$$