Will Governments Fix What Markets Cannot? The Positive Political Economy of Regulation in Markets with Overconfident Consumers

Patrick L. Warren and Daniel H. Wood*

October 2010

Abstract
A standard theoretical and empirical claim in the behavioral industrial organization literature is that market forces may not always eliminate the negative welfare consequences of biased consumer behavior, even in competitive markets. In this literature, government regulations are often proposed that would succeed in raising social welfare in the face of consumer biases and could therefore serve as substitutes for market forces. It is not obvious, however, that a self-governing group of citizen-consumers would choose to enact this welfare-improving policy. We show that in the case of markets for goods with add-ons (Gabaix and Laibson, 2006) the instances in which markets or government will fail to correct for consumer biases (in our model, consumer overconfidence about add-on services such as overdraft fees) are similar. In the same cases in which markets are inefficient due to the prevalence of biased consumers, voters will not demand efficiency-enhancing regulations. Consumer biases have two effects: they produce deadweight losses, and they redistribute income from biased consumers to less-biased consumers. The distributional consequences often both prevent equilibrium competition by firms from enhancing efficiency and prevent equilibrium policy choices by citizens from regulating away inefficient trade.

JEL Classification:

Keywords: shrouded attributes, myopia, market failure, government failure, behavioral economics, behavior public policy, behavioral industrial organization.

*Department of Economics, Clemson University. patrick.lee.warren@gmail.com, dwood2@clemson.edu.
1 Introduction

If imperfectly rational consumers exist, their choices can make themselves and – through equilibrium effects – others worse off than they would be if the consumers were more fully rational. It has long been argued that market competition could limit the impact of biased consumer choices if firms compete for the rents available from exploiting biases. A variety of recent papers, however, provide theories and evidence that markets sometimes fail to restore efficiency when consumers are unaware of their biases (Gabaix and Laibson, 2006; Grubb, 2009; Heidhues and Koszegi, 2009). Within the context of these models, appropriate policy responses will be able to increase social welfare. As a result, most behavioral public policy has focused on normatively optimal regulations to help biased consumers (Gabaix and Laibson (2006, section IV.B); Bar-Gill (2006, section IV); Heidhues and Koszegi (2009, section II.B)).

Less attention has been paid to the positive question of when normatively optimal policies are likely to be adopted. We take existing explanations of when competition fails to “fix” consumer biases as given. We ask whether a self-governing set of citizens with policies available that could completely eliminate problems caused by consumer biases will adopt those policies. There are cases where the citizens will, but only if the ability of the firms to commit are particularly circumscribed. Otherwise, direct governmental intervention runs afoul of the same problems that keep market forces from correcting consumer biases in the first place. If consumer overconfidence leads to “too much” demand in the market for goods, it may also lead to “too little” demand in the market for regulation.

To illustrate this idea, we consider a market for add-ons in the presence of overconfident consumers.1 Our model shares many similarities with that of Gabaix and Laibson (2006).2 Firms sell firm-specific base goods or services and also add-ons for those goods. An add-on in our model is a supplemental good (such as water in a hotel room’s mini-fridge) or service (such as overdrafts on bank accounts) that some consumers will require in order to properly enjoy the consumption of the base good.

---

1 See Moore and Healy (2008) for a survey of the psychology literature on overconfidence.
2 Like Gabaix and Laibson, our model will explain high markups on add-ons by positing the existence of boundedly rational consumers. There are several models of add-on pricing with perfectly rational consumers, most of which view add-on pricing as a method of price discrimination. One stream simply assumes an inability to commit to add-on prices (Borenstein, MacKie-Mason, and Netz 1995; Farrell and Klemperer 2005; Klemperer 1987). Another, including Ellison (2005), views add-ons as made possible by relatively large exogenous search costs. Our purpose is not to contrast rational and behavioral explanations, but rather to show that if add-ons are driven by factors related to bounded rationality, we should be particularly pessimistic about regulation as a solution.
In our model, the market for the base good is perfectly competitive, but firms are monopolists in the add-on market. Furthermore, firms have limited abilities to commit to add-on prices *ex ante*.

Consumers take into account the potential future lock-in when they decide whether to buy the base good or not. Consumers at greater risk of needing the add-on will only buy the base good when it is relatively cheap. With this market structure, firms compete away their add-on profits by pricing the base good below marginal cost. High-risk purchasers cross-subsidize low-risk purchasers – low-risk purchasers benefit from the subsidy to base goods while seldom having to pay the inflated price for add-ons. For example, using a credit card with rewards is better than using cash, if the balance is paid off every month. The money the credit card pays out in rewards come from other consumers who failed to pay off their balances.

Overconfidence that one is at low risk of needing the add-on can lead to inefficient consumption, as overconfident, risky consumers who should not consume the good at the market price do so. In contrast, high-risk consumers who are rational would never consume when they are made worse off by consumption. This overconsumption by the overconfident further subsidizes the base good, and can lead to consumption by all types, even when the cost of production exceeds the value of the good to the consumers. Furthermore, even the rational high-risk consumer may prefer to continue to patronize the firms with high add-on prices, in order to take advantage of the overconfident.

We examine the implications of overconfidence for these kinds of markets, but our main contribution is to link this model to the positive question of what regulation will be demanded by consumers. We ask what policy might occur as the outcome of some political interaction of the agents in our model (Meltzer and Richard (1981)). We find that capping the add-on price at the cost of production is efficient, but in the presence of overconfident consumers, an inefficiently high cap will often be favored by some or all of the citizens. The intuition here is that no one thinks that they are among the overconfident, and they want a high cap to take advantage of the overconfident. When regulation does occur, it is not to prevent biased individuals from being exploited, but rather because of the side-effects of this exploitation harm rational individuals more than they benefit from the exploitation. Widespread consumer overconfidence thus leads to policy failure as well as market failure.

While our focus is overconfidence about goods with add-ons, this connection between market failure and policy failure is more general. The fundamental reason that
either markets or government would want to “cure” biased consumers is that their biases result in deadweight loss. However, these losses will also be associated with net transfers among consumers (or to other interested parties such as firms) and so the distributional results of a cure simultaneously limit palliation by markets and by government.

Our results will be interesting to two constituencies: the first and obvious one is behavioral economists interested in public policy. The second one is more skeptical economists who are nonetheless curious about the logical implications of the current stream of behavioral industrial organization research. In order to cater to both groups, we delay until the last section the major application of our model, a consideration of how the modern credit card industry and its regulation, which we anticipate to be of greater interest to the first group. Aside from that, our organization is standard.

2 Related Literature

Several behavioral industrial organization papers analyze how responses by firms affect whether consumers with particular decision-making biases are saved from the biases’ harmful effects.\(^3\) In a 2009 review, DellaVigna notes that the welfare implications of nonstandard preferences often depend on whether consumers have rational expectations or nonrational expectations about those preferences. For example time-inconsistent consumers who are aware of their time-inconsistency are willing to pay for commitment devices, and so will in equilibrium attain welfare-maximizing contracts (but see Gottlieb (2008) for when commitment devices are not possible). With consumers who are unaware of their time-inconsistency, though, firms often make higher profits by taking advantage of them by offering tempting contracts (DellaVigna, 2009). This distinction is true of this paper; if all consumers in our model were Bayesian, there would be no social welfare losses in equilibrium.

Consumer behavior in our model combines aspects of the behavior in Grubb (2009) and Gabaix and Laibson (2006). Grubb looks at three-part tariffs offered to consumers who overestimate the precision of their demand forecasts. He applies that analysis to explaining cellular phone contracts. Consumers in our model make mistakes in their demand forecasts for add-on goods that could be microfounded as overconfidence about the precision of their forecasts.

There is no natural distinction between base goods and add-ons in the markets

\(^3\)See Ellison (2006) for a survey of behavioral IO.
that Grubb analyzes, in contrast to our model and that of Gabaix and Laibson (2006). In Gabaix and Laibson’s base model, the authors consider a market for a good with an add-on that is required by all consumers, some of whom are completely unaware of the add-on. In our model needing an add-on is a risk that all consumers face, but some consumers will not need one. All consumers are aware of the add-on, but some have incorrect beliefs about their risk, replacing the total unawareness of Gabaix and Laibson.

Another way overconfidence sometimes manifests in behavioral economics models is that some agents do not correctly forecast their future self-control problems. Hyperbolic discounting is a form of time-inconsistent preferences in which people weight future utility less than present utility (Laibson, 1997; O’Donoghue and Rabin, 2006). Many applications of hyperbolic discounting are based upon “naïve” hyperbolic discounters who believe that in the future they will have more self-control than they do in the present. Partial naïveté can explain pathological behavior where employees take months or years to sign up for 401k plans (O’Donoghue and Rabin, 1999).

Heidhues and Koszegi examine the implications of naïve hyperbolic discounting for credit markets. In their model, firms offer inexpensive baseline credit, but with large penalties for delaying repayment. Naïve consumers overborrow and end up suffering these penalties, leading to welfare losses (Heidhues and Koszegi, 2009). These choices are similar to those that overconfident consumers make in our model when the net utility value of the good is low. Eliaz and Spiegler (2006) is an earlier contract-theoretic model in this vein. In that paper, agents are time inconsistent and vary in their naïveté. The principal offers contracts that allow commitment and also contracts that exploit consumer irrationality.

While the arguments in this paper are couched in terms of overconfident consumer behavior, the same logic applies to bounded rationality more generally in markets with aftermarket services or add-ons, as the least rational people will generally end up subsidizing more rational consumers. For example, Stango and Zinman (2009a) finds that some consumers take much more expensive loans if APR information is not available. They attribute this to some consumers failing to understand exponential growth. If these consumers are unaware of the higher add-on premiums they pay and are also unaware that they are unaware, they will not demand public policy interventions that would alleviate the harm they suffer as a result of their mistakes.

In our model, the policy response citizens decide upon is a cap to add-on prices. In reality, there are other forms of policy response to the sorts of problems brought
on by overconfidence. First, governments could attempt to “debias” consumers, for example by requiring firms to educate their customers. Bertrand and Morse (2009), for instance, finds 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 (see, for example, Hastings and Weinstein (2008) or Mastrobuoni and Weinberg (2009)). Furthermore, educational interventions will likely fall prey to the same under-demand problems that we identify.

Much of the economics literature on policy responses proposes informational solutions, but a second form of response seems at least as popular with advocacy groups and politicians, namely to directly intervene in the market. This regulation may be quite heavy-handed, like outlawing certain transactions or capping certain fees. Alternatively, it may be relatively light, such as a “nudge” towards certain options via strategically legislating a default option (Camerer et al., 2003; Thaler and Sunstein, 2009). We focus on the most direct form of response in this paper, modeling regulation as a simple price cap, but our results would extend to more “libertarian” brands of regulation. Price caps largely act as instruments for closing down the ex-post monopolistic add-on market, so citizens will have similar decisions about other policy instruments that would limit after-market profits.\(^4\)

We are aware of only one paper that looks at how biases in consumer markets will play out in the corresponding regulatory “markets”. Ed Glaeser is also pessimistic about behavioral public policy, but for different reasons than us. He argues that consumers have weaker incentives to choose well when voting for politicians than when buying products or services and that politicians may be easier to mislead than consumers in general (Glaeser, 2006). Our paper is a complement to Glaeser’s work, since our focus is not on regulatory implementation but on the more basic demand for regulation.

\(^4\)Obviously heavy-handed regulation typically has more distortionary effects than mandating a default would, but we abstract away from these implementation problems. If voters are skeptical about the efficiency of government regulation, it will just make regulation even more unlikely to fix biases when markets do not.
3 Base Model: Selling to Rational Consumers Without Price Commitment

3.1 Primitives

Payoffs and Production We consider a market with a base good and an add-on. There is a unit mass of consumers who receive a utility of 1 for consuming the base good, but may be subject to an event which reduces this value by \( k \), which event we will term “needing the add-on”. The loss \( k \) may be the pure effect of the event or the consumer’s cheapest way of responding to the event other than purchasing the add-on. Consumers have two alternatives to accepting this utility-reducing event. First, if the event does occur they can consume an add-on good, which completely removes the effect of the event, restoring the lost \( k \). Second, before the occurrence of the event is realized, they can engage in an avoidance behavior at cost \( E \). If they do so, the event will never occur.

A set of perfectly competitive producers produce the base good at constant marginal cost \( C \). There is free entry into the base good market.\(^5\) The add-on good is costless to produce, and it is specific to the base good produced by any given producer. So while there is perfect competition in the base good market, producers are monopolists in the market for their add-on.

Actions and Timing We divide market behavior into five periods. In the first period, firms set their prices simultaneously, where firm \( i \) chooses price \( p_{B,i} \) for the base good and \( p_{A,i} \) for the add-on. In the second period, consumers observe these prices and choose whether to buy the base good and from whom. A firm pays the cost of production if and only if a consumer buys its base good.

In the third period, consumers who have purchased the base good decide whether to engage in avoidance behavior. In the fourth period, firms may have the opportunity to reset their add-on price. This opportunity to renege arises with probability \( 1 - r \). Here, \( r \) measures the firms’ ability to commit to the add-on price. Inability to commit may arise due to consumer inattention, class problems with incomplete contracts, or the inability of consumers to discriminate perfectly between a binding promise and cheap talk or marketing.\(^6\) In the last period, consumers learn whether or not they

\(^5\)A competitive base good market is not necessary for our results. Appendix A shows that they would still occur if the seller of the base good were a monopolist.

\(^6\)Many of the restrictions placed on credit cards by the Credit CARD Act support the idea
are subject to the welfare shock and then decide whether or not they want to buy the add-on.

**Types and Information** There are three sorts of consumers who can differ both with respect to their risk of experiencing a utility-diminishing event and with respect to their beliefs about this risk. High risk consumers will suffer this event with probability $\alpha_h$ and low risk consumers will do so with probability $\alpha_l < \alpha_h$. High risk consumers make up a fraction $\beta_h$ of all consumers, while low-risk consumers make up the remainder.

In addition, a fraction $\gamma_o$ of the high-risk consumers are overconfident and believe incorrectly that they are low risk. They are perfectly certain about this belief but incorrect. Consumers (believe they) know their own type during period 1, but firms can never observe consumers’ types. The distribution of types is common knowledge, including the presence of the overconfident types, as are the firms’ costs.

All firms have identical costs. We focus on symmetric equilibria where all firms offer the same contract or set of contracts.

Finally, when making their original purchase decisions, consumers have to form beliefs about the final add-on price. All consumers are Bayesian in their beliefs about prices.

### 3.2 Relating the Model to Markets

To help fix ideas about the sorts of markets this model may help us understand, consider an interpretation of the model’s primitives in the following markets.

**Checking Accounts** The base good is the basic service associated with having a checking account, including electronic transfers, writing checks, access via ATMs, and so forth. One utility-reducing event which may occur is writing a check which goes over the balance in the account, leading to no payment on that check and the associated embarrassment, penalties, and credit score consequences or else a last-minute

that consumers were either unable to anticipate all contingencies fully or did not believe seemingly beneficent offers were legally binding.

7 The non-Bayesian nature of the overconfident consumers’ beliefs is necessary for their presence to affect the market equilibrium. If consumers were Bayesian, so that upon learning they were “low risk”, they actually inferred that they had risk $\alpha_h$ with probability $\gamma/(1 - \beta)$, then they would avoid the add-on (for low $E$) or not consume (for $K > 1$). With Bayesians, this would be exactly equivalent to increasing $\alpha_l$ and reducing $\beta$. Overconfident people generate add-on revenue that no consumer thinks they themselves are paying for and all consumers benefit from.
costly liquidation of assets to cover the check. The add-on is overdraft protection, which guarantees payment for small overages, but is accompanied with a fee payment to the bank. Avoidance is taking precautions like monitoring deposits and withdrawals on your account carefully and frequently to be sure never to write check that might bounce. There are a number of other events and fees which apply to checking accounts as well: using an ATM that is out of network, exceeding the maximum number of withdrawals in a month, and dropping below a minimum balance.

**Credit Cards**  The base good consists of the convenience associated with using a credit card to make purchases and the ability to finance the purchase of large items or unexpected expenses. The event is a sudden cash flow problem making it difficult to pay the minimum monthly payment, to the extent that scraping together sufficient funds in time would involve a significant utility cost of $k$. The add-on consists of skipping the payment and paying the accompanying fees and APR penalties of $p_A$. Avoidance could represent more disciplined use of the credit card or the opportunity costs of maintaining a “rainy day” fund to cover the minimum payment in case of cash flow shock.

**Cell Phone Contracts**  The base good consists of an annual contract with certain restrictions and a limit on monthly minutes. The event is a sudden and unexpected increase in the consumer’s demand for cellular telephone usage in a given month. He could forgo this extra usage or use a substitute service, at cost $k$, or he could pay $p_{A,i}$ to buy the additional minutes. In order to avoid this problem, he could have carefully monitored his minutes throughout the month, making only essential calls and using a land line as much as possible. In this case, the demand spike would not have cleared the monthly allocation of minutes, but all that avoidance behavior would have been costly $(E)$.  

**Room Service**  The base good is a hotel room. The shock is tiredness, so that it is very uncomfortable to go out to dinner. The add-on is room service or eating from a mini-fridge. Avoidance is packing something to drink or eat for dinner.
3.3 Perfect Bayesian Equilibrium with Exogenous Add-on Pricing

A useful starting point in understanding this market and how regulation might affect it is to first consider the case in which the add-on price is exogenously specified. This exercise is useful for three reasons. First, in a market where the price is set by regulation, this case will describe the complete economic equilibrium. Second, when there is no ability to commit \( r=0 \) all firms will set \( p_A = k \) when given the opportunity to renege. Consumers will perfectly foresee this, and act accordingly in the first period. Finally, any contract offered in the general game with complete commitment ability must fit all the conditions derived here, plus the additional condition that firms must not want to deviate from the contract or set of contracts offered.

If the add-on price is exogenously specified, there is a unique symmetric perfect Bayesian equilibrium in the market described above. Its structure turns on the costs of avoidance and the gains from trade in the market. Let \( G \equiv 1 - C \) represent the potential gains from trade in the market for the base good. Then, the following proposition describes its key features of this equilibrium.

**Lemma 1.** Assume that the add-on price is exogenously fixed at \( p_A \leq k \), and the game is otherwise exactly as described above. There is a unique symmetric perfect Bayesian equilibrium of the game described above.

1. For consumers who believe they are type-i, there is a threshold gain from trade \( G_i(E) \), where they forgo the base good in equilibrium if \( G < G_i(E) \).

2. \( G_l(E) \leq G_h(E) \) for all \( E \).

3. Conditional on buying the base good, consumers who believe they are type-i avoid if \( E < p_A \alpha_i \).

4. If \( G \geq G_l(E) \) so some consumers buy the base good, producers all charge \( p_B = C - p_A D_A \) and make zero profits, where \( D_A \) is the expected proportion of base good purchasers who purchase the add-on as well.

Figure 1 represents the equilibria described by the above proposition. Proofs of this lemma and other formal claims are given in appendix B.

\[ \beta_h = \gamma_a = 1/2, \quad \alpha_l = 1/5, \quad \text{and} \quad \alpha_h = 4/5. \]
Figure 1: Types of Equilibrium by Gains from Trade and Cost of Avoidance. In region 1 no consumers purchase the base good or the add-on. In region 2, all consumers purchase the base good and avoid the add-on. In region 3, low-risk consumers purchase both the base good and the add-on. In regions 4 and 5, all low-risk consumers purchase both and some high-risk consumers purchase the base good and avoid the add-on. Finally in region 6 all consumers purchase both the base good and the add-on. All regions except for 6 are not socially optimal relative to charging $C$ for the base good and 0 for the add-on.
The primary mechanic that drives behavior in the sort of market we analyze is the way add-on revenues subsidize base-good pricing. Since the add-on is costless, the revenues generated by add-on sales add directly to the bottom line, and so are competed away as lower base-good prices. Customers differentially face the burden of the add-on, but share in common the base-good discount. If all consumers were equally at risk for the ex-post shock, the net price would adjust completely back to marginal-cost pricing, but since they are differentially vulnerable, there are distributional consequences. In particular, if the expected demand for the add-on among those buying the base good is \( D_A \), a consumer of type \( i \) who buys the good and the add-on receives a net subsidy of \( p_A(D_A - \alpha_i) \), while a consumer who buys the base-good and avoids receives a net subsidy of \( p_A D_A - E \).

For low avoidance costs \( (E < \alpha_LP_A) \), all consumer types will avoid, and so there is no demand for the add-on and the base good will be priced at marginal cost. For this range of parameters, there is no differential pricing by risk type, since no consumers purchase the add-on. Everyone gets a “subsidy” of \(-E\). Furthermore, overconfidence plays no role, since all types act identically. Despite this marginal cost pricing, there are two inefficiencies in the market. First, the avoidance behavior wastes \( E \) in effort. Second, the foreseen avoidance waste leads consumers to under-consume – that is, not purchase the base good despite some positive gains from trade (region 1). In particular, the trade is consummated only if the gains from trade exceed the cost of avoidance (region 2).

For moderate avoidance costs \( (\alpha_LP_A < E < \alpha_HP_A) \), conditional on purchasing the base good, those believing themselves to be high-risk will avoid, while those believing themselves to be low-risk will not. The positive expected demand for the add-on, coupled with the zero-profit condition in the base good market, leads to the base good being priced below marginal cost. Producers expect to make up for the losses on the base good with add-on revenues. If the gain from trade is high enough, some (region 4) or all (region 5) high-risk consumers purchase the base good and avoid the add-on. The subsidy plus the gains from trade are sufficient to cover their avoidance costs. But if the gains from trade are too small, only those believing themselves to be low-risk will be willing to buy the base good (region 3). Combined with the fact that the high-risk will avoid, this means that any equilibrium in this range with positive gains from trade will either involve under-consumption, avoidance, or both.

In this range of avoidance costs, overconfidence begins to affect the equilibrium outcomes. Imagine for a moment that there were no overconfident consumers. Then
when gains from trade are small, only the low risk buy the base good and the expected conditional demand for the add-on is $\alpha_l$. These expected revenues will be factored into the price of the base-good, it will be discounted by exactly $p_A \alpha_l$. Taken together, then, the expected total lifetime cost of purchasing the good for a low-risk consumer is exactly $C$, the marginal cost of production. Absent consumer heterogeneity, the extra add-on costs are exactly offset by a lowering of the base-good price. Adding the overconfident back in increases the expected add-on demand (since they are actually high risk), further subsidizing the base-good price, driving the net price for low risk below the cost of production. Thus, we get low-risk consumers buying the good, even when there are small losses from trade. Finally, if the subsidization from the overconfident is large enough, relative to the costs of avoidance, it will even lure in the high risk, who buy the highly subsidized base good and then avoid the costly add-on. The overconfident are subsidizing the rational.

When the cost of avoidance becomes very high ($E > \alpha_h p_A$), even the high-risk consumers no longer avoid. As in the middle range, high-risk consumers will buy the base good only if the gains from trade are high enough. Note that they do so despite the fact that they are thereby subsidizing the low-risk consumers, since they will pay for the add-on with a relatively high probability. If the gains from trade get too small, the high-risk will exit the market, and the equilibrium will look exactly like it did for the moderate costs of avoidance. Furthermore, unlike the moderate-cost case, the cutoff gain from trade is independent of the fraction overconfident, since no type will avoid regardless of their belief about their type.

Welfare For every parameter configuration, the producers make zero profits, but the welfare of consumers varies significantly. The following lemma summarizes the welfare consequences brought about by the add-on problem. For our purposes, an “efficient” outcome is one that produces the same social welfare (total consumer utility plus firm profits) as would be produced by firms selling at prices $p_B = C$ and $p_A = 0$. Consumer utility is their experienced utility, measured ex post rather than relative to their ex ante beliefs. When the add-on price is above zero, equilibrium outcomes can be inefficient in three different ways: 1) “avoidance”: some or all consumers expend non-productive effort $E$ to avoid needing the add-on, 2) “underconsumption”: some or all consumers do not purchase the base good despite $G > 0$, or 3) “overconsumption”: some or all consumers purchase the base good despite $G < 0$.

**Lemma 2.** 1. When there are positive gains from trade, the market with add-ons
Figure 2: Forms of Deadweight Loss in Equilibrium by Gains from Trade and Cost of Avoidance. Diagonal lines indicate regions where underconsumption occurs, shaded area indicates regions where overconsumption occurs, and dots indicate regions where avoidance occurs.

is efficient if and only if avoidance is very costly \((E > \alpha_h p_A)\) and the gains from trade are large \((G > G_h(E))\). Otherwise the equilibrium in the game involves avoidance, under-consumption, or both.

2. When there are small negative gains from trade \((G_l(E) < G < 0)\), the market is efficient if and only if the cost of avoidance is very low \((E < \alpha_l p_A)\). Otherwise, there is overconsumption.

3. As the fraction of the high risk who are overconfident increase

   \((a)\) The maximum equilibrium losses from overconsumption increase.

   \((b)\) The maximum equilibrium losses from underconsumption decline.

   \((c)\) Fewer citizens avoid, but they avoid for lower gains from trade.

Figure 2 shows the regions in \((G, E)\)-space where each form of inefficiency occurs.
4 Endogenous Pricing and Regulation

Lemma 2 categorized the possible sources of inefficiency if firms charge some exogenous add-on price $p_A$. Unless the costs of avoidance are quite high and the gains from trade are large, we should worry about the efficiency of any market that only sells goods with significant add-on prices.

This section investigates how competitive firms will set their add-on prices, and how citizens might try to affect these choices through regulation. We represent the possibility that competition in the market will result in efficiency by considering what happens when firms can set add-on prices prices freely and commit to them in the first period ($r=1$). This is a particularly strong form of competition, as firms essentially lose their ex-post market power.\(^9\)

If markets are unable to solve the problem, regulation still could. We take a positive approach to regulation, in which the regulation imposed is determined by the will of the citizens participating in the market. Our measure of whether positive political regulation will succeed is whether (some or all) consumers prefer regulations on the maximum add-on price that would restore efficiency to the market. The presence of overconfident consumers strongly affects whether robust markets or political regulation will produce efficient equilibria.

We find that when markets would not guarantee efficiency, the political process always fails to deliver efficiency as well. This occurs because of similar dynamics in both cases. Markets do not function efficiently when consumers would rather get a heavily subsidized base good, taking advantage of overconfident consumers, than purchase goods at marginal cost. The political process in turn fails when voters would rather get subsidies due to overconfident consumers than mandate an efficient pricing scheme.

4.1 Perfect Competition and The Add-on Price

Consider first whether perfect competition in the market would lead to efficient add-on pricing.

**Proposition 1.** Assume firms commit to an add-on price in period 1 perfectly ($r=1$). When there are no overconfident consumers ($\gamma_o = 0$), all equilibria are efficient and include $P_b = C$ and $P_a = 0$.

However,

\(^9\)In Ellison (2005) this would be concomitant to search costs going to zero.
1. when \( \gamma_o > 0 \), for any \( G \) such that \( G_l(k) < G < 0 \) (with \( G_l(k) \) from lemma 1) and \( E > (\frac{G}{G_l(k)}) k \alpha_l \), any equilibrium includes overconsumption.

2. there is a cutoff

\[
E^c = k(\beta_h \gamma_o \alpha_h + (1 - \beta_h) \alpha_l) \left( \frac{\gamma_o \alpha_h - \alpha_l}{\gamma \alpha_h - \gamma \alpha_l} \right),
\]

such that for any \( E < E^c \) and any \( G > 0 \), any equilibrium includes avoidance.

It is not particularly surprising that when consumers are rational, perfectly competitive markets are efficient. The only potential market imperfection is the ex-post monopoly power, and that inefficiency is competed away with add-on price commitment in the first period. The more surprising results, although consistent with the recent literature on the industrial organization of markets with imperfectly rational consumers, concern the inability of competition to solve the problems when the overconfident make up a large fraction of society.

In particular, competition has no significant impact on overconsumption. When there are negative gains from trade, consumption occurs because the overconfident pay too much for the add-on, in expectation. A firm committing to a small add-on price would attract no customers, since he will need to charge more the consumer’s value for the base good in order to make non-negative profits, so there are no profitable deviations away from the equilibrium described in lemma 1 with \( k \geq p_A > 0 \).

The ability of competition to solve problems other than overconsumption hangs on whether there is a separating equilibrium in which low-risk consumers choose a “riskier” offer with \( p_B < c \) and \( p_A > 0 \) and high-risk consumers choose a less risky offer, for example marginal-cost pricing of \( p_B = C, p_A = 0 \). Intuitively, if there is no separating equilibrium, it must be because the high risk prefer pooling. Because marginal cost contracts are feasible, the only reason the high risk would prefer to pool would be if they then avoided the add-on. However, because marginal cost contracts are possible, high-risk consumers will always be willing to buy the base good, so commitment does eliminate all underconsumption.

4.2 Equilibrium Regulation

Return now to the model with exogenous \( p_A \) in order to consider the scope for regulation. Without commitment, \( p_A = k \), and the market will always suffer from deadweight loss, unless the gains from trade and the costs of avoidance are both
very large relative to the price of the add-on (region 6 in Figure 1). The most straightforward regulatory solution to this dilemma, then, would be to limit the price of the add-on, thereby eliminating the ill-effects of the firms’ ex-post market power. In particular, a regulation which fixes the price that producers can charge for the add-on at 0 would guarantee efficiency, since neither consumer type will avoid, and the price of the base good would be priced at marginal cost.

There are several ways we could think about this price limitation, which are substantively different but would be equivalent in the model. The first and most straightforward is that the limitation is a simple price cap, limiting by law the price that the firms are allowed to collect from consumers for providing the add-on, with or without renegotiation. An alternative interpretation is that the regulation is some limit on market behavior that improves the consumer’s outside option, $-k$. For example, the recent Credit CARD act requires a longer notification period before any major change in terms, allowing consumers a greater opportunity to search for alternatives to paying the add-on price. For either interpretation, we assume this regulation could be costlessly and perfectly enforced. This will give the greatest scope for useful regulation.

A government price regulation could improve the overall efficiency in the market, but it also has distributional impacts. When there are overconfident consumers, these distributional consequences will sometimes limit support for a price control.

**Proposition 2.** Without overconfident consumers ($\gamma_o = 0$), each consumer’s preferred add-on price regulation is efficient, as is any weighted average of their preferred price caps.

However,

1. **(overconsumption)** When $\gamma_o > 0$, for any $G_l(k) < G < 0$ and $E > (\frac{G_l}{G_l(k)})k\alpha_l$, no consumer strictly prefers an efficient add-on price and low-risk consumers strictly prefer an inefficient add-on price to any efficient price.

2. **(underconsumption)** When $\gamma_o > 0$, there is a cutoff

$$G_l^* = \frac{k(1 - \beta_h)\gamma_o(\alpha_h - \alpha_l)}{\beta_h\gamma_o + 1 - \beta_h}$$

such that for any $G$ that satisfies $0 < G < G_l^*$ and any

$$E > k(\frac{G}{G_l^*})\left[\frac{\beta_h\gamma_o\alpha_h + (1 - \beta_h)(1 - \gamma_o)\alpha_l}{\beta_h\gamma_o + (1 - \beta_h)(1 - \gamma_o)}\right].$$
low-risk consumers strictly prefer an inefficient add-on price to any efficient price.

3. (avoidance) When $\gamma_o > \frac{\alpha_l}{\alpha_h}$, there is a cutoff

$$E_h^* = k(\beta_h \gamma_o \alpha_h + (1 - \beta_h) \alpha_l)$$

and when $\gamma_o > \frac{\alpha_l}{\alpha_h}(2 - \frac{\alpha_l}{\alpha_h})$ there is a cutoff

$$E_l^* = k\alpha_h \left( \frac{\gamma_o \alpha_h - \alpha_l}{\alpha_h - \alpha_l} \right)$$

such that for any $E < E_i^*$ and any $G > 0$, type-i consumers strictly prefer an inefficient add-on price to any efficient price.

With no overconfident consumers, capping add-on prices will always improve the welfare of the high-risk types. In fact, the efficient cap of $p_A = 0$ will always maximize the welfare of the high-risk types, since paying for the add-ons always results in a transfer from the high-risk consumers to the low-risk consumers, and avoidance is pure loss.

The low-risk types, by contrast, always prefer a higher cap. The high-risk types’ high probability of being “caught” by the add-on subsidizes the low-risk types’ consumption of the base good. This subsidy is dependent on two requirements: there must be a positive price for the add-on and the high-risk types must pay that price. These requirements will limit the low-risk type’s desire to set a high price-cap. But these two requirements are exactly a guarantee against underconsumption and against avoidance, so the low-risk will prefer the largest cap which eliminates the potential market imperfections.

In the presence of overconfident consumers, efficient regulation involves a trade-off. By correcting the market imperfection, it increases total consumer surplus in the market, but at the same time limits the scope for rational consumers to profit from the mistakes made by the overconfident. When the overconfident are particular numerous or profoundly overconfident, rational consumers will prefer to keep the market inefficient and take advantage of the overconfident. The overconfident, believing themselves to be rational, will support this as well.

For the high risk, if the subsidy on the base good generated by selling to the overconfident is larger than the cost of avoidance, they will prefer an inefficiently large price regulation to maximize this subsidy. Of course, this strategy depends on
a low cost of avoidance, and enough overconfident to fund this transfer. Furthermore if the gains from trade are negative \((G < 0)\), the high-risk would not even buy the base good with the efficient price regulation, but they weakly prefer all the inefficient regulations in that case.

The optimal price regulation for the low risk is slightly more complicated. They will choose it either to maximize the cross-subsidization by the high risk, as in the case without overconfident consumers, or to maximize the cross-subsidization by the overconfident. The value to the low risk of cross-subsidization by the high risk depends on the degree to which they need to reduce the add-on price to convince high risk consumers to buy and not avoid. As the cost of avoidance declines, lower and lower prices would be necessary to induce efficient action among the high-risk. If there are sufficient numbers of overconfident consumers, there is a point at which the low-risk consumer will give up on enticing the high-risk consumers and simply charge a high add-on price and enjoy the subsidy from the overconfident. Small gains from trade have the same flavor, where an efficient price would need to be quite low to entice the high-risk consumers to buy the base good. As the efficient price approaches zero, the low risk are getting no subsidy at all from the high risk, so they would prefer instead keep the add-on price high and collect a subsidy from the overconfident.

Finally, if the gains from trade are negative, no efficient price would lead to consumption by any consumers. The unregulated equilibrium, by contrast, includes overconsumption when the loss from trade is not too great. The low risk strictly prefer that over no consumption, so they would never support an efficient regulation in this region.

### 4.3 Competition versus Regulation

Figure 3 represents the parameter combinations for which regulation and competition correct the potential inefficiencies in this market. A cursory inspection reveals that these two figures share a number of important similarities. First, there are large regions of the parameter space for which neither response suffices to correct the inefficiencies, particularly when the cost of avoidance is low or the gains from trade and negative.

When the high-risk consumers control regulation, the range of parameters for which regulation is effective is smaller than the range for which competition solves the problem. Competition is effective when firms can appeal to each type separately. One attractive contract for the high risk is a low add-on price coupled with a relatively
(a) Regulation by High Risk and Low Risk. Vertical shading indicates region in which low risk strictly prefer inefficient equilibria. Horizontal shading indicates region in which high risk strictly prefer inefficient equilibria.

(b) Competition. Light shading indicates region where increased competition does not induce efficient equilibria.

Figure 3: When Competition Fixes Inefficiency versus When Government Fixes Inefficiency.
high base good price. We argued in the previous section that the high risk’s most preferred efficient regulation has approximately this form. The high risk therefore find the committed package attractive in all cases where they support efficient regulation. This overlap means that if the high risk control regulation, we should never expect it to be able to solve problems that robust competition cannot.

When the low-risk consumers control regulation, the effective ranges for commitment and regulation still overlap to a great degree, but there are two exceptions. First, when the gains from trade are small and positive, competition might be able to induce efficient consumption (if the cost of avoidance is large enough), but regulation never will. Second, when there are few overconfident consumers, the low risk may implement efficient regulation even when competition would not induce it. Conversely, if there are many overconfident consumers, the low risk may impose inefficient regulation even when competition would lead to efficiency. The \( \gamma_0 \) that equalizes \( E_l^* \) and \( E_h^* \) in Proposition 2 is the point at which competition and high-risk consumers lead to the same equilibrium.

Note that both of these interventions are completely ineffective in dealing with overconsumption. Quite the contrary, one perverse effect that both these interventions can have is to increase the range of parameters for which overconsumption occurs. When the cost of avoidance is very low (\( E < p_A \alpha_l \)), the base model does not suffer from overconsumption, since all consumer types avoid (even the overconfident), so there is no inefficient subsidization. With the possibility of competition or regulation, the firms or regulator can lower the add-on price to \( E/\alpha_l \) and thereby induce the overconfident to no longer avoid, opening up the subsidization dynamic and extending overconsumption into low costs of avoidance.

The comparisons in this section revealed two circumstances under which we might expect the government to outperform the market. First, if robust competition by way of add-on price commitment is not possible for technical reasons, there is certainly the possibility of effective regulation for relatively high avoidance costs when the high risk control policy, and high avoidance costs and gains from trade when the low risk control policy. Second, even if commitment is feasible, regulation may still have an edge when there are relatively few overconfident consumers, the gains from trade are large, and the regulatory authority is controlled by the high-risk. In all other circumstances, the robust market will outperform equilibrium regulation or they will be equally effective.
4.4 Triggers of Efficient Regulation

The induced preferences for regulation outlined above can be used to predict or explain when regulations capping the add-on price in a market will appear. In particular, we can predict what sorts of changes in the economy would induce the citizens to support the imposition of an efficient cap. These shifts come in three varieties: changes in populations, changes in avoidance costs, and changes in risks.

Regardless of which consumer type controls policy, the range of parameters for which efficient regulation will be imposed grows as the fraction overconfident in the population decreases (the product $\gamma_o\beta_h$ declines) The intuition for this result is that the size of the subsidy from the overconfident is proportional to their share of the population, so it becomes less attractive to obtain this subsidy at the cost of avoidance (for the high risk) and the cost of forgone subsidies from the high risk (for the low risk).

A similar result holds for ex-ante avoidance costs ($E$), where, independent of who controls policy, a push for efficient regulation may be triggered as the cost of avoidance increases. For the high risk, this result follows because they take advantage of the overconfident through buying the subsidized base good and avoiding the add-on. For the low risk, the intuition is different. For them, an efficient price cap is attractive because it entices the high risk to buy the base good and risk the add-on, subsidizing the price of the base good. If avoidance becomes more difficult, a higher price cap will be sufficient to entice the high risk, increasing the returns from efficient regulation.

The mirror of this result occurs for the changes in the costs born by a consumer who refuses to buy the add-on ($k$), which acts as a sort of exogenous “cap” on the add-on price. Regardless of who controls policy, a push for efficient regulation can be spurred by a reduction in this cap, since size of the subsidy from the overconfident in the absence of regulation is proportional to that price.

Changes in riskiness have more subtle effects on the demands for regulation. Regardless of who controls policy, a reduction in the riskiness of the high risk ($\alpha_h$ declines) will make regulation more likely. Efficient regulation is opposed in order to take advantage of the overconfident, and this becomes less attractive for all types as the overconfident become less risky.

Changes in the riskiness of the low risk will affect different consumers’ differently. The low-risk consumers’ desire for efficient regulation will increase as the low risk become riskier ($\alpha_l$ increases). As the riskiness of the low risk increases by one unit, their expected equilibrium payoff will fall by the price of the add-on. But the add-
on is priced lower under efficient regulation, so their payoff is less affected in that instance, leading a marginal low-risk consumer to prefer regulation as his riskiness increases.

When the high risk control policy, however, an increase in the riskiness of the low risk can actually spur a push for deregulation. As the riskiness of the low risk increases by one unit, the expected equilibrium payoff of the high risk will rise by the price of the add-on. In direct contrast, then, a large add-on price is attractive, since there is more to gain from buying the discounted base good and avoiding the add-on, leading the high risk to support inefficient regulation.

Taken together, these considerations offer a number of plausible explanations for the timing of the recent Credit CARD act enactment. First, it may have been triggered by an increase in the riskiness of the low-risk types. If the mortgage crisis and associated increase in unemployment led to a systemic increase in the riskiness of otherwise responsible credit card users, this increased risk may have led those consumers to support a law limiting credit card add-on fees. Second, it may have been triggered by a growth in alternative lending institutions such as payday lenders or car title loans. If these sources of cash represented an cheaper alternative to missing a credit card payment (lower $k$), then their arrival would reduce the size of the subsidy all types received from the overconfident, perhaps leading consumers to instead prefer efficient regulation.

5 Regulation and the Credit Card Industry

As an illustration of the ideas contained in our model, consider the market for credit cards. The base good is the capacity to make purchases without carrying around cash or checks, combined with a line of credit. The base good price might be quite small: a fixed interest rate to be paid on balances transferred onto the card at the start or an annual fee. Credit card add-ons come in two types. The first are significant fees triggered by contingencies. These add-on fees can amount to hundreds of dollars per year for some consumers (Stango and Zinman, 2009b). Although the fees appear formally in the contractual terms, many consumers are unaware of their exact conditions in practice. Furthermore, until recent regulation, add-on fees could be changed with little customer notification, or with notification cloaked in extremely difficult language and small print. Our model represents these ambiguous and shifting fee structures by an imperfect ability to commit.
A second form of add-on is interest charges on balances. Many consumers seem overconfident about their ability to repay revolving credit in a timely manner. 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. (2005) find that when choosing between an offer with a fee and a lower interest rate or no fee and a higher interest rate, many borrowers took the offer with no fee despite the fee offer being less costly ex post.

Heidhues and Koszegi (2009) analyze the equilibrium contracts that would be offered to credit card and sub-prime mortgage borrowers in a market similar to our paper’s, in which a subset of consumers are time-inconsistent and overconfident about the interest charges they will pay. They find that the equilibrium contract has front-loaded terms, which lead to consumer default along with the associated fees. They conclude that there is scope for welfare-improving regulatory interventions that limit the size of penalties. Heidhues and Koszegi, however, do not explore the demand for these regulations among the consumers.

Consumers exhibit great heterogeneity in how these add-ons impact them. Stango and Zinman (2009b) calculate month-to-month correlations in fees and interest costs consumers suffer. The 75th percentile most consistent in paying high fees have a month-to-month correlation of 0.56 and the 75th percentile most consistent in paying high interest costs have a month-to-month correlation of 0.49. A minority of consumers pays the majority of the add-ons of each type.

The credit card industry fits our model’s assumptions well. There is high variation among consumers in the change that they will occur these add-on costs. Many consumers exhibit overconfidence about their repayment ability (and presumably about their chance of occurring fees). Firms are unable to commit fully to the price of these add-ons. Anecdotal reports of exploitative fees and interest traps are also common, suggesting many consumers are unhappy with some of the industry’s practices.

Recently U.S. Federal regulations have placed restrictions on what credit card companies can charge for these add-ons. Pricing come into existence. The Credit CARD Act of 2009 prohibits many of the practices of the credit card industry that critics claimed were exploitative. Many of the changes amounted to limiting add-on fees and interest charges. Rates can no longer increase during the first year of having the card. Very explicit notification for APR changes are mandated and the degree of advanced notification must be at least 45 days. Universal default – the practice of
increasing card users’ interest rates based on their payment records with unrelated accounts, such as utilities or other credit cards – is banned.

In our model, mandated cuts to add-on prices lead to an increase in base good prices, and we see some evidence of this in the credit card context. In a recent survey of direct mail credit card offers, the fraction of cards requiring an annual fee has increased to the highest level in a decade. Base interest rates have also markedly increased. The implementation of the new regulations has corresponded to a decline in rewards programs.

This trade-off suggests that for some consumers, those who enjoyed rewards programs and were at low risk for fees (or believed they were at low risk), the Credit CARD Act was harmful. Historically, the credit card industry opposed the sorts of reforms implemented in the Act by warning that they would have to raise prices (increase fees or reduce rewards) in response. This tactic may explain the relative lack of regulation until this point, and the relative weakness of the regulation that was imposed.

Yet some regulation did occur. The recession of December 2007 to June 2009 seriously increased the risks of credit card use for responsible consumers. In our model terms, the probability that low-risk individuals needed the add-on increased. As analyzed in section 4.4, this changing risk demographic makes cross-subsidization by the overconfident less appealing, and if the shift is large enough, it can produce a political will for efficient regulation. Note however that the intent of the regulation is not to prevent exploitation per se, but rather to limit the side effects of the redistributive “technology” used to benefit the rational at the cost of the overconfident.

The specifics of the Credit CARD Act buttresses this interpretation. Appendix C details the major provisions of the Act. With the exception of provisions placing limits on upfront prices for cards for borrowers with subprime credit and limiting the availability of credit cards for people under age 21, none of the provisions directly targets consumers who would overestimate their ability to repay the revolving credit. Many of the provisions make interest rates and fees more predictable, changes which benefit rational consumers making optimal plans more than overconfident consumers. While several add-on fees are targeted (such as double-cycle billing or shifting due dates), again these are add-ons which impose heavy costs on non-overconfident consumers as well.

6 Conclusions and Implications

The primary conclusion we draw from the analysis in this paper is a pessimistic one for the prospects for effective imposition of regulation. We build a model of overconfident consumers in a market with add-ons, and show that consumer overconfidence leads to persistent deadweight loss due to unproductive avoidance, underconsumption, or overconsumption.

We first show that more competition, by way of add-on price commitment, solves all the market imperfections if all consumers are perfectly rational. With overconfident consumers, commitment only solves some of the problems, leaving overconsumption and avoidance for a subset of the parameters. The intuition here is that producers would use commitment to commit not to soak the risky consumers and thereby entice them to consume and not avoid. But in accepting this offer the risky consumers forego the opportunity to take advantage of the overconfident, so it becomes less attractive as there are more overconfident consumers.

Since more competition does not solve everything, and may not even be feasible, we in turn consider regulating the market for add-ons. If all consumers were perfectly rational, they would all prefer an efficient price cap on add-ons (although they would disagree on the exact form it would take). With enough overconfident consumers, however, there is an wide range of parameters for which there will no demand from any citizens for efficient regulation. The intuition is straightforward. In the absence of regulation, the pie is smaller but a large part of the overconfident’s share of the pie is transferred to the rational. Since no one thinks they are overconfident, there is little reason to support a policy to protect them. Furthermore, the parameter space for which there is a demand for efficient regulation is actually more restrictive than the space for which perfect competition is effective.

When regulation does occur, it is because the business methods used to collect rents from overconfident consumers are also harmful to rational consumers. In section 5 we argue that recent credit card regulations seem to match this story better than an alternative story where regulation is directly intended to help biased consumers.

We conclude that, even if competition in the market is quite robust, there may remain important negative welfare consequences brought about by boundedly rational agents. Even so, the prospects for effective regulation in these circumstances are especially poor. The same forces which make avoidance or overconsumption hard problems for the market to resolve make them a hard problem for the government to resolve. Only if the market is particularly hampered (e.g., by an inability to commit,
as in our base model), might we be guardedly optimistic about the government’s ability to improve on market outcomes. But, note, this is not an instance of the government intervening to correct for imperfections brought about by boundedly rational agents, but rather a much more classical exercise of authority, where the government is regulating in order to complete ineffectual markets.
7 References


A Monopoly

Consider an alteration of the base model where, instead of assuming perfect competition in the market for the base good, we assume a monopolist produces the base good. Everything else in the model is unchanged. Consumer’s decisions, conditional on prices are unaffected by this change, and, in the absence of commitment, so is pricing of the add-on. All that changes is the pricing of the base good, where the monopolist will choose the base-good price to maximize profits. His primary decision, then, boils down to one pursuing one of two strategies:

1. Sell to all consumers, and sacrifice some consumer surplus to the low-risk types.

2. Sell to only those believing themselves to be low-risk consumers, taking all of their consumer surplus plus an additional transfer from the overconfident, but sacrificing the gains from trade from the high-risk.

Which strategy they pursue will depend on the gains from trade and the fraction of each type in the population. The following proposition summarizes the equilibrium. The proofs for all the propositions follow very closely the proofs for the perfectly competitive case. They are available by request.

Proposition 3. Let $G \equiv 1 - C$ represent the gains from trade in the market for the base good. A perfect Bayesian equilibrium of the game with monopoly producer exists. In all such equilibria

1. For consumers who believe they are type-i, there is a threshold gain from trade $G_i(E)$, where the monopolist sells to consumers who believe they are of type $i$ if $G \geq G_i(E)$

2. $G_l(E) \leq G_h(E)$ for all $E$.

3. Conditional on buying the base good, consumers who believe they are type-i avoid if $E < k\alpha_i$.

4. The monopolist charges $p_A = k$ and $p_B$ to extract all the consumer surplus from the highest-risk type who buys.

The welfare consequences are very similar to the perfect competition case, although the exact cutoffs differ.

Proposition 4. 1. Unless the cost of avoidance is very high ($E \geq k\alpha_h$) and the gains from trade are very large ($G \geq k(\alpha_h - \alpha_l)[\frac{1}{\beta_h(1-\gamma_o)} - 1]$), there will be underconsumption or avoidance or both whenever there are potential gains from trade.

2. If the cost of avoidance is moderate or high ($E > \alpha_l k$) and the gains from trade are small and negative ($G > k\frac{\beta_h \gamma_o (\alpha_l - \alpha_h)}{\beta_h \gamma_o + 1 - \beta_h}$), there will be overconsumption.
The scope for commitment to solve the market inefficiencies in the monopoly context is actually better than in the perfectly competitive context. With two consumer types of the sort in this simple economy, the monopolist is able to costlessly separate them using second-degree price discrimination. The perfectly competitive market does something like this in the case of high avoidance costs, but at small costs of avoidance, the contract targeted at the low-risk and overconfident consumers also attracts the high-risk who subsequently avoid. The monopolist prevents this by setting a relatively high base-good price \((G - k\alpha_l)\) and a high add-on price \(k\) on one contract, and \((G, 0)\) on the other.

**Proposition 5.** If a monopolist can commit to add-on prices in period 1, underconsumption and avoidance disappear, but overconsumption remains.

Finally, if consumers can regulate both prices, all consumer types will want to hold the monopolized firm to zero expected profits. Conditional on an add-on price cap, consumers will unanimously agree to cap the base good price at the same price charged by the perfectly competitive firm. Given the agreement on the base price, preferences over the the add-on price cap will exactly coincide with those in the base model.

**Proposition 6.** If consumers can regulate both the add-on and base good price, but non-negative expected profits are required to ensure production, the regulated equilibria with a monopoly producer include the same prices and consumption choices as that with perfectly competitive producers.

## B Proofs

### B.1 Proof of Lemma 1

**Low avoidance costs:** If all consumers prefer to avoid the add-on \((E < \alpha_l < \alpha_h)\), then the demand for the add-on is 0. Firms do not profit from the aftermarket and so set the base price \(p_B = C\). If \(G > G_L(E) = G_H(E) = E\), all consumers will purchase the base good.

**Medium avoidance costs:** If only consumers who believe they are low risk buy the base good, then demand for the add-on is

\[
D_A = \frac{\gamma \beta_h \alpha_h + (1 - \beta_h) \alpha_l}{\gamma \beta_h + 1 - \beta_h}
\]
the average risks of the low risk and overconfident (high risk) consumers, weighted by their frequencies. Given this demand, low risk consumers will want to buy if

\[ 1 - (C - kD_A) - k\alpha_l > 0 \]

\[ G > G_L(E) = -k\left(\frac{\gamma o\beta_h(\alpha_h - \alpha_l)}{\gamma\beta_h + 1 - \beta_h}\right) \]

and high risk consumers will not want to buy if

\[ 1 - (C - kD_A) - E < 0 \]

\[ G > G_H(E) = E - k\left(\frac{\gamma\beta_h\alpha_h + (1 - \beta_h)\alpha_l}{\gamma\beta_h + 1 - \beta_h}\right) \]

If all consumers buy the base good then demand for the add-on is

\[ D_A = \gamma\beta_h\alpha_h + (1 - \beta_h)\alpha_l. \]

High risk consumers will want to purchase the add-on as long as

\[ 1 - (C - k[\gamma\beta_h\alpha_h - (1 - \beta_h)\alpha_l]) - E > 0 \]

\[ G > E - k[\gamma\beta_h\alpha_h - (1 - \beta_h)\alpha_l]. \]

For an intermediate range of high-risk consumer surplus

\[ k(\gamma\beta_h\alpha_h + (1 - \beta_h)\alpha_l) > G - E < k\left(\frac{\gamma\beta_h\alpha_h + (1 - \beta_h)\alpha_l}{\gamma\beta_h + 1 - \beta_h}\right), \]

some but not all high-risk consumers purchase the base good. This is because high risk purchasers increase the first period price in equilibrium by avoiding and paying nothing in the third period, but getting part of the first period rebate. A fraction \( z \) of high risk types will buy, where \( z \) is such that high risk types are indifferent. Demand for the add-on is then

\[ D_A = \frac{\gamma\beta_h\alpha_h + (1 - \beta_h)\alpha_l}{\gamma\beta_h + 1 - \beta_h + \beta(1 - \gamma_o)z}. \]

Zero profit in period 1 implies \( p_B = C - kD_A \) and so high risk types are indifferent when

\[ 1 - C + k\left(\frac{\gamma\beta_h\alpha_h + (1 - \beta_h)\alpha_l}{\gamma\beta_h + 1 - \beta_h + \beta(1 - \gamma_o)z}\right) - E = 0 \]

\[ z = \frac{k(\gamma\beta_h\alpha_h + (1 - \beta_h)\alpha_l) + (1 - C - E)(\gamma\beta_h + 1 - \beta_h)}{(C + E - 1)(1 - \gamma_o)\beta_h} \]

**High avoidance costs:** If only low risk consumers purchase the base good, demand for the add-on good is \( \frac{\gamma\beta_h\alpha_h + (1 - \beta_h)\alpha_l}{\gamma\beta_h + 1 - \beta_h} \). For the associated first period discount, low
risk consumers prefer to purchase as long as
\[ 1 - C + k \left( \frac{\gamma \beta \alpha + (1 - \beta) \alpha}{\gamma \beta + 1 - \beta} \right) - k \alpha > 0 \]
\[ G > G_L(E) = -\frac{k \gamma \beta (\alpha - \alpha)}{\gamma \beta + 1 - \beta} \]

High risk consumers will prefer not to purchase the base good as long as
\[ 1 - C + k \left( \frac{\gamma \beta \alpha + (1 - \beta) \alpha}{\gamma \beta + 1 - \beta} \right) - k \alpha < 0 \]
\[ G < \frac{k(1 - \beta)(\alpha - \alpha)}{\gamma \beta + 1 - \beta} \]

If all consumers purchase the base good, demand for the add-on good is \( \pi \). High risk consumers prefer to purchase because of the associated first period discount when
\[ 1 - C + k \left[ \gamma \beta \alpha + (1 - \beta) \alpha \right] - k \alpha > 0 \]
\[ G > G_H(E) = k(1 - \beta)(\alpha - \alpha). \]

Unlike the medium avoidance cost case, there is no region in which high risk purchasers mix between purchasing and not. In the region
\[ k(1 - \beta)(\alpha - \alpha) < G < \frac{k(1 - \beta)(\alpha - \alpha)}{\gamma \beta + 1 - \beta} \]
if high risk purchasers do purchase they prefer to purchase, but if they did not purchase they would prefer not to purchase. However, the former will be the only equilibrium decision. In this range of gains from trade, if all firms were selling only to low risk consumers, \( p_B = C - \frac{\gamma \beta \alpha + (1 - \beta) \alpha}{\gamma \beta + 1 - \beta} \) any firm would prefer to deviate and sell to all consumers at a profit. High risk types would buy if
\[ 1 - p_B - k \alpha > 0 \]
\[ p_B < 1 - k \alpha. \]

At that price the deviating firm’s profits would be
\[ 1 - k \alpha + k \left[ \beta \alpha + (1 - \beta) \alpha \right] - C = \]
\[ G - k(1 - \beta)(\alpha - \alpha) > \]
\[ k(1 - \beta)(\alpha - \alpha) - k(1 - \beta)(\alpha - \alpha) = 0 \]
profits. Hence, if an equilibrium where high-risk consumers purchase the base good exists, it is the unique equilibrium.
B.2 Proof of Lemma 2

If $G < G_l(E)$ or $G < 0$ and $E < k\alpha_l$, no one purchases the base good. If $G > G_h(E)$ and $E > k\alpha_h$ everyone purchases the base good and no avoids. Both of these cases are efficient.

**Avoidance:** If high-risk consumers purchase, they avoid as long as $E < k\alpha_h$.

**Underconsumption:** If $G < G_h(E)$, no high-risk consumers purchase the base good.

**Overconsumption:** If $G_l(E) < G < 0$, there is overconsumption.

**Comparative statics:** As $\gamma_o$ increases, $G_l(E)$ and $G_h(E)$ fall. Underconsumption is less frequent and overconsumption is more frequent. Increasing $\gamma_o$ reduces avoidance because only high-risk consumers avoid. □

B.3 Proof of Proposition 1

Let $\tilde{\alpha}$ be the expected demand for the add-on conditional on only consumers believing themselves to be low risk buying the base good and not avoiding. Similarly, let $\bar{\alpha}$ be the expected demand for the add-on when all consumers buy the base good and do not avoid.

$$\tilde{\alpha} = \frac{\beta_h \gamma_o \alpha_h + (1 - \beta_h) \alpha_l}{\beta_h \gamma_o + 1 - \beta_h} \quad \bar{\alpha} = \beta_h \alpha_h + (1 - \beta_h) \alpha_l$$

**No overconfident consumers:** Equilibrium offers to each type of consumer make zero profit. Type $i$ consumers receive an offer $p_{Bi} = c - \alpha_i p_{Ai}$ where $0 \leq p_{Ai} \leq k$. Each must prefer their own contract to the other’s:

$$U_l(p_{Bi}, p_{Ai}) \geq U_l(p_{Bh}, p_{Ah})$$
$$G \geq G + \alpha_h p_{Ah} - \alpha_l p_{Ah}$$
$$G \geq G + p_{Ah}(\alpha_h - \alpha_l)$$

so $p_{Bh} = c$ and $p_{Ah} = 0$. Consumers are always offered a marginal cost contract, so there is no inefficiency.

**Avoidance:** Any efficient equilibrium with commitment and $\gamma_o > 0$ will be a separating equilibrium. If it were a pooling equilibrium, it would have a base good price $p_B = c - \tilde{\alpha} p_A$ due to the zero-profit condition for firms. In this case, if $p_A > 0$, a profitable deviation for a firm is to offer $p_{Bd} = p_B + \alpha_h \epsilon - \delta$ and $p_{Ad} = p_A - \epsilon$ for small $\epsilon, \delta > 0$. $p_{Ad}$ and $p_{Bd}$ attracts the high risk only and makes positive profits. Pooling contracts with $p_A > 0$ are ruled out, but since $p_A$ must be non-negative, $p_A = 0$ must still be considered. In this case, a profitable deviation is $p_{Bd} = c - \tilde{\alpha} p_A + \delta$ and
Only low risk consumers are attracted to this deviation and it produces positive profits for the firm.

Any efficient equilibrium must be a separating one. Competition for low risk consumers leads to their receiving the highest \( p_A \) offer compatible with high risk consumers not taking that offer. Low risk consumers perceive their utility to be increasing at rate \( dU_l/dp_A = (\tilde{\alpha} - \alpha_L) \), but in fact \( d\pi/dp_A = \tilde{\alpha} \), so firms bid up the add-on price of the offer they make to the low risk. Having high risk consumers not take the low risk offer and avoid limits the add-on price. \( p_B = c - \tilde{\alpha}p_A \), and so \( U_h = G + \tilde{\alpha}p_A - E \) if they avoid. Their own contract will give them \( G \), so \( p_{Al} = \min\{k, E/\tilde{\alpha}\} \).

The high risk contract is not pinned down to a point. It must be the case that

\[
U_l(p_{Bl}, p_{Ah}) \geq U_l(p_{Bh}, p_{Ah})
\]

\[
G + p_{Ah}(\tilde{\alpha} - \alpha_l) \geq G + p_{Ah}(\alpha_h - \alpha_l)
\]

\[
p_{Ah}\gamma_o(\alpha_h - \alpha_l) \geq p_{Ah}(\alpha_h - \alpha_l)
\]

so \( 0 \leq p_{Ah} \leq p_{Ah}\gamma_o/(\beta_h\gamma_o + 1 - \beta) \) and \( p_{Bh} = c - \alpha_h p_{Ah} \).

Since this separating equilibrium is the only efficient equilibrium, if firms want to deviate from any separating contract, then there is no efficient equilibrium.

Consider possible deviations when \( p_{Al} = E/\tilde{\alpha} \). An increase in \( p_A \) will require lowering \( p_B \) if it will attract any consumers, and hence will lead to high risk consumers taking that contract and avoiding the higher add-on price. For the deviation to attract the low risk, their utility must increase. It will increase the most if \( p_{Ad} = k \). This gives a condition on \( p_{Bd} \):

\[
U_l(p_{Bl}, p_{Ad}) \geq U_l(p_{Bh}, p_{Ah})
\]

\[
1 - p_{Bd} - \alpha_l k \geq G + \frac{E(\tilde{\alpha} - \alpha_l)}{\tilde{\alpha}}
\]

so

\[
p_{Bd} \leq C - \alpha_l k - \frac{E(\tilde{\alpha} - \alpha_l)}{\tilde{\alpha}}
\]

The deviating firm makes profits

\[
\pi = p_{Bd} + k(\beta_h\gamma_o\alpha_h + (1 - \beta_h) - \gamma\alpha_l)
\]

\[
< -\alpha_l k - \frac{E(\tilde{\alpha} - \alpha_l)}{\tilde{\alpha}} + k(\beta_h\gamma_o\alpha_h + (1 - \beta_h))
\]

\[
= k\beta_h(\gamma_o\alpha_h - \alpha_l) - \frac{E\beta_h\gamma_o(\alpha_h - \alpha_l)}{\beta_h\gamma_o\alpha_h + (1 - \beta_h)\alpha_l}
\]

This will be negative if \( E > k(\beta_h\gamma_o\alpha_h + (1 - \beta_h)(\gamma\alpha_h - \alpha_l))/(\gamma\alpha_h - \gamma\alpha_l) \).
Overconsumption: The low risk believe they get utility \(1 - p_B - p_A\alpha_l\) from a contract with prices \(p_B\) and \(p_A\). The most profitable contract for firms will be \(p_A = k\) because \(\pi = p_B + p_A\frac{\gamma_h\alpha_h + (1 - \beta_h)\alpha_l}{\gamma_h + 1 - \beta_h} - C\) increases in \(p_A\) faster than low risk utility falls in \(p_A\). At that price, the maximum \(p_B\) possible is \(p_B = 1 - k\alpha_l\). Profits from this price combination are

\[
p_B + k\frac{\gamma_h\alpha_h + (1 - \beta_h)\alpha_l}{\gamma_h + 1 - \beta_h} - C =
\]

\[
G - k\alpha_l + \frac{\gamma_h\alpha_h + (1 - \beta_h)\alpha_l}{\gamma_h + 1 - \beta_h} =
\]

\[
G + k\frac{\gamma_o\beta_h(\alpha_h - \alpha_l)}{\gamma_h + 1 - \beta_h}.
\]

At \(G < G\), no contract that consumers will accept is profitable for firms. If \(G < G < 0\), there may be a profitable contract low risk consumers will accept. For there to be one, because gains from trade on the base good are negative, firms must profit off the add-ons. Only if \(E > p_A\alpha_l\) will the overconfident and low risk purchase add-ons. When \(p_A \geq E/\alpha_l\), profit is negative for all acceptable contracts satisfying this inequality if \(E < (\frac{G}{\alpha_l})k\alpha_l\).

\[\Box\]

B.4 Proof of Proposition 2

The equilibrium payoff of a type-i player is buying the base good is given by \(G + p_A D_A(G, E, p_A) - \min\{E, p_A\alpha_l\}\), otherwise he gets 0.

No overconfident consumers: Assume, first, that \(\gamma_o = 0\). There are three cases: both consumer types buy the add-on, so \(D_A = \alpha < \alpha_h\); neither type buys the add-on, so \(D_A = 0\); or only the low-risk buy the add-on (and some high-risk may buy the base good and avoid), so \(D_A \leq \alpha_l < E/p_A < \alpha_h\). In all three cases the net transfer to the high-risk consumers is negative, while if he caps at \(p_A = 0\), the net transfer is zero, so he will always strictly prefer such cap when \(G > 0\).

For the low-risk, the net subsidy is positive if and only if the high-risk buy the base good and do not avoid, but is otherwise increasing in \(p_A\). So the low risk prefer the highest \(p_A\) which is consistent with both. Such a price is efficient by construction, and exists if \(G > 0\), since a small \(p_A = \epsilon\) works in that instance.

If \(G \leq 0\), the no consumers ever make positive utility, so they are both indifferent among price caps.

Overconfident consumers: Now, assume \(\gamma_o > 0\). Nothing changes from the three cases above except the last case. Now when only those believing themselves to be low-risk buy the add-on, but everyone buys the base good, the expected add-on demand is \(D_A = \beta_h\gamma_o\alpha_h + (1 - \beta_h)\alpha_l\). When only those believing themselves to be low-risk buy the add-on or the base good, expected add-on demand is \(D_A = \frac{\beta_h\gamma_o\alpha_h + (1 - \beta_h)\alpha_l}{\beta_h\gamma_o + (1 - \beta_h)}\).
**Overconsumption:** When $G < G < 0$, and $E > k\alpha_l$ there is equilibrium over-consumption in the absence of a price cap, by proposition 1. An efficient price cap would eliminate this consumption, making the low-risk strictly worse off, so they would not support it. When $E < k\alpha_l$, and $G < 0$ there would be no consumption without a price cap, but a cap at $p_A = \frac{E}{\alpha_l}$ would induce those believing themselves to be low-risk to purchase the base good. It follows directly from reapplication of proposition 1 with $k' = \frac{E}{\alpha_l}$ and the forgoing that the low-risk would support this cap if $0 > G > \frac{E}{k\alpha_l}$. Rearrangement gives the expression in the proposition.

**Underconsumption:** Consider $G^*_l = k(1 - \beta_h)\gamma_o(\alpha_h - \alpha_l)\beta_h\gamma_o + 1 - \beta_h$. There is some $E$ (given in lemma 1) such that the high risk do not consume for gains from trade $G^*_l$. That $E$ is $k\gamma_o\alpha_h + (1 - \beta_h)(1 - \gamma_o)\alpha_l + 1 - \beta_h$.

When $0 < G < G^*_l$ and $E \geq k\left[\frac{\gamma_o\alpha_h + (1 - \beta_h)(1 - \gamma_o)\alpha_l}{\beta_h\gamma_o + (1 - \beta_h)}\right]$, only those believing themselves to be low-risk are buying the base good in the unregulated equilibrium (The threshold $E$ is that which makes the first potential high-risk purchaser just indifferent, the point on line separating regions 3 and 4 at $G^*_l$ in Figure 1). The equilibrium payoff to the low risk is $G + k\left[\frac{\gamma_o\alpha_h + (1 - \beta_h)(1 - \gamma_o)\alpha_l}{\beta_h\gamma_o + (1 - \beta_h)}\right] - E$. Direct comparison shows that the low risk prefer the inefficient payoff if $G < G^*_l$.

When $0 < G < G^*_l$ and $E < k\left[\frac{\gamma_o\alpha_h + (1 - \beta_h)(1 - \gamma_o)\alpha_l}{\beta_h\gamma_o + (1 - \beta_h)}\right]$, some high risk may prefer to purchase the base good and avoid the add-on, reducing the attractiveness of the unregulated equilibrium to the low risk. Lowering the add-on price has no effect on the base good price, since it drives out the high risk who are avoiding (some of the high risk avoid and some do not purchase here), and thus makes the low risk better off. In particular, a cap $z$ such that $E = z\left[\frac{\gamma_o\alpha_h + (1 - \beta_h)(1 - \gamma_o)\alpha_l}{\beta_h\gamma_o + (1 - \beta_h)}\right]$, will lead to no consumption of the base good by the high risk. This reduces the problem exactly to the prior case, but $k = z$.

**Avoidance:** Finally, when $\gamma_o > \alpha_l/\alpha_h$, consider first the preferences of the high risk. Of all efficient price caps, they prefer a cap at 0, giving them a payoff of $G$. Absent a price cap, when $E > k\alpha_l$, they could potentially do better by buying the base good and avoiding the add-on. This nets them a payoff of $G + k[\beta_h\gamma_o\alpha_h + (1 - \beta_h)\alpha_l] - E$, so they strictly prefer no cap in accordance with the proposition. If $E < k\alpha_l$ a price
cap at \( z = \frac{E}{\alpha_l} \) would be necessary to ensure the low risk and overconfident do not avoid. Following the argument above, the high risk prefer that cap to the efficient one if \( \frac{E}{\alpha_l}[\beta_h \gamma_o \alpha_h + (1 - \beta_h)\alpha_l] > E \). But that inequality is always satisfied when \( \gamma_o > \frac{\alpha_l}{\alpha_h} \).

Now consider the preferences of the low risk. Absent a cap, when \( k\alpha_h > E > k\alpha_l \) the low risk receive an equilibrium payoff of no less than \( G + k\beta_h (\gamma_o \alpha_h - \alpha_l) \), the payoff when the high risk all buy the base good and avoid the add-on. An efficient price cap must induce that high risk to not avoid, so it can be no greater than \( \frac{E}{\alpha_h} \). The low risk’s payoff with an efficient cap, therefore, is no greater than \( G + \frac{E}{\alpha_h} \beta_h (\alpha_h - \alpha_l) \), and \( E^*_l \) equalizes these payoffs. When \( E < k\alpha_l \), even an inefficient policy must cap that add-on price in order to induce the overconfident to buy the base good. Since higher add-on prices are always more attractive to the low risk, the most attractive inefficient cap is \( \frac{E}{\alpha_l} \). Replacing for that in the first expression above, the inefficient policy is preferred to the efficient if \( \gamma_o \frac{\alpha_h}{\alpha_l} - 1 > 1 - \frac{\alpha_l}{\alpha_h} \), which any \( \gamma_o \) above that given in the proposition satisfies.

\[ \square \]

### C Provisions of the Credit CARD Act

#### C.1 Interest Rates

1. Retroactive interest rate hikes banned without
   - (a) end of a teaser rate, or
   - (b) explicit renegotiation with cardholder, or
   - (c) indexed interest rate, or
   - (d) 60 days+ default, and then return to base rate after 6 months on time payment.

2. No rate increases in first year unless explicitly so advertised up front as a teaser rate. then none in first 6 months.

3. No universal default (raising rates on cards because of a default on other, un-related account, like utility bill).

#### C.2 Double-Cycle Billing

4. Bans method of calculating balance, which allowed finance charges from a second period to apply to the first period too.

#### C.3 Gift Cards

5. No expiration for at least 5 years unless clearly advertised as such; No dormancy fee for less than 12 months inactivity.
C.4 Timely Payments
6. Bill must be mailed at least 21 days before due date.

C.5 Subprime Cards
7. No up-front fees of more than 25% of available credit limit in first year.

C.6 Credit Allocations
8. Payments above minimum must be applied first to highest rate debts.

C.7 Due Dates
9. Due date must be the same each month. If they fall on a weekend, payments on next business day must be accepted without penalty.
10. If creditor changes address or payment procedures, cannot charge fees for 60 days after.

C.8 Over-Limit Fees
11. Must be option to opt-in/out of over-limit fees. Otherwise, excessive purchases will be rejected with no fee. Only one over-limit fee per cycle.

C.9 Disclosures
12. 45 day notice for significant changes in terms, including interest rates and fees.
13. Notification of the right to close and that closing will not require an immediate payment of balance.
14. Disclosure of how long it will take to pay off debt and what the total interest charges will be at minimum payment
15. Disclose of monthly payment to pay off debt in 36 months.

C.10 Youth and Credit
17. No cards for people under 21 without either adult co-signers or proof of independent means.
18. No pre-screened offers for those under 21.
19. No freebies for signing up near a college campus or college-sponsored events.
20. College-marketing contracts must be disclosed and filed with the Fed.