Elicitation and Corrective Taxation

Akhil Rao and Brennan McConnell

LATEST DRAFT: May 12, 2018

Abstract

Marginal contributions to observable aggregate stocks are often unobservable in games with negative stock externalities, making optimal corrective taxation a difficult endeavor. We propose a new class of mechanism, the *elicited tax*, for such settings. The elicited tax uses an observable aggregate to elicit private information about marginal contributions, and a scored tax to penalize reports which are inconsistent with the observable aggregate and other reports. We define a notion of strict propriety for elicited taxes, show that under perfect competition reports are Nash equilibria if and only if they are consistent with the observable aggregate, and that strictly proper elicited taxes ensure socially optimal output and externality production. We focus on a particular strictly proper elicited tax, the *Brier-Pigou* tax, which combines a modified Brier scoring rule with a Pigouvian tax. Numerical experiments highlight three properties of the Brier-Pigou tax: (1) the tax can achieve the social planner's welfare when firms are perfectly competitive; (2) the tax achieves close to the planner's welfare even when firms are perfectly collusive; and (3) the proportion of consistent lies which are risk-dominated by truthful reporting for each firm depends on the number of firms providing reports and that firm's true marginal externality.

1 Introduction

The idea of soliciting reports from agents to levy a tax on them or assess their compliance with a regulation is not new. For example, income taxes administered by the Internal Revenue Service involve asking individuals about their earnings and then levying a tax based on their report. Vehicle emissions standards involve asking manufacturers to avoid gaming measurement devices to assess their compliance, and potentially levy a fine if they are found to be noncompliant. These are examples of *elicited taxes*: taxes whose magnitude must be elicited from the very agents being taxed. However, despite their widespread use, elicited taxes have not been explicitly considered in the elicitation or optimal taxation literatures. To our knowledge this paper is the first to formally explore this idea in connection with externality-correcting mechanisms.

Standard externality-correcting devices assume either that the regulator can observe each agent's marginal contributions to the aggregate externality or that the regulator knows the function mapping observable outcomes to externality generation. We consider a setting where perfectly competitive firms' marginal contributions to a stock externality depend on private information which is not observable to the regulator. We describe a class of mechanisms with which a regulator can use observable information and reports by firms to encourage socially optimal behavior.

Such problems are recognized in dispersed and nonpoint pollution control problems such as orbital debris generation, where the debris generated by a firm's satellite may not be attributable to the satellite; vehicle emissions reporting, where the regulator cannot observe each vehicle's pollution generation in real-time over the full range of driving behaviors; and with agricultural runoff into streams, where the regulator is unable to attribute levels of pollutants in the water to specific farms.

Questions of uncertainty and incentives in group externality optimization have a long history in environmental economics and game theory. Segerson Segerson (1988) built on earlier work about incentives and moral hazard in teams (Groves (1973), Holmstrom (1982)) to show a mechanism which could ensure efficient production without knowledge of individual firms' marginal abatement activities. However, that mechanism requires firms to have consistent and correct conjectures about the distribution of ambient levels, that the regulator knows how marginal abatement activities by individual firms change the probability of observing an ambient level, and that the damages are linear in abatement activities. While the mechanism can be efficient in the long run, it is not budget-balanced, as firms are charged more than the total amount of damages caused by the externality. Though this allows the possibility of a double-dividend, it would seem to complicate the political economy of implementing such a mechanism.

Since the mechanism proposed in Segerson (1988), researchers have provided mechanisms which can account for nonlinearities in the damage function Hansen (1973), explored the use of group and individual penalties to enforce policy measures Kritikos (2004), and studied the importance of targeting to cost-effectiveness Baerenklau (2002). We explore a slightly different setting in this paper. First, we suppose that an aggregate stock is measured with confidence, and that the regulator knows how marginal production activities by firms generate the externality up to a single-dimensional parameter known only to individual firms. Then we ask, "how can we incentivize firms to behave optimally for society without levying more than the socially optimal level of taxes?" Strictly proper elicited taxes are one answer to this question.

The model presented here can be related to the one in Segerson (1988) as follows. Suppose we (the regulator) observed ambient pollution levels, and had a sufficiently-good model of the physical processes involved to turn ambient measurements into point estimates of the aggregate stock generated by a group of firms in a specific period. Suppose also that we knew the mapping from abatement activities to marginal pollution contributions for a given firm up to a single dimensional parameter known only to the firm. Under these conditions, strictly proper elicited taxes are able to achieve socially efficient aggregate outcomes and marginal pollution production (i.e., abatement). Further, in every Nash equilibrium, the mechanism is exactly budgetbalanced. We offer simulation evidence that deviations from truthfulness tend to be small and follow some regular patterns, suggesting that further work in this area may be able to ensure truthful reporting as well as efficiency.

A simple mechanism that could achieve efficient production and efficient total taxation without elicitation would be to charge each of n firms $1/n$ of the full amount of the measured externality. Then marginal incentives for production are preserved, and the amount of tax revenue generated covers the cost of the externality but no more. However, this mechanism leaves the firms' marginal abatement activities unknown to the regulator, and doesn't allow firms to report the truth. The presence of n in the marginal taxes may also distort entry decisions, and less efficient firms are likely produce at the expense of more efficient firms in the long run.

One advantage of elicited taxes over this simple mechanism is that elicited taxes allow marginal taxes to vary. We show that in equilibrium, output and externality production will be socially optimal, and simulation results suggest that truthful reporting will be a locally risk-dominant Nash equilibrium. Perfect collusion in reporting will involve inconsistent reports and socially inefficient output and externality production, but simulation results suggest that the welfare loss from perfect reporting collusion is bounded and small.

In the following section, we present the model setting. In section 3, we prove some properties of elicited taxes in general and explore the Brier-Pigou tax using simulations. Section 4 suggests some directions for future research, and Section 5 concludes.

2 Model

2.1 Setting

Suppose there are n perfectly competitive firms in the market producing an output good and an externality. Let x_i be firm i's production of the good, p the world price of the good, $c_i(x_i)$ be i's cost to producing x_i units of the good, and ϵ_i be i's privately known externality generation parameter. Pollution is produced by the production process according to the function $f(x_i, \epsilon_i)$, which is strictly convex and increasing in x_i and ϵ_i. For simplicity, we assume that each unit of the externality causes a constant amount of social damage and that the cost of a single unit of damage is normalized to 1, so that the total social cost of D units of externality is simply D.

Definition 1. *The aggregate observable externality is the sum of individual firm externalities generated, i.e.*

$$D = \sum_{i=1}^{n} f(x_i, \epsilon_i)$$

Definition 2. *The social welfare function is*

$$W = \sum_{i=1}^{n} [px_i - c_i(x_i) - f(x_i, \epsilon_i)].$$

The social planner chooses output to maximize social welfare, solving

$$\max_{\{x_i\}_{i=1}^{n}} \quad \sum_{i=1}^{n} [px_i - c_i(x_i) - f(x_i, \epsilon_i)]$$

The first order conditions to the planner's problem show that the optimal output decision satisfies

$$x_i^{planner} \quad : \quad p - c_i^x(x_i) - f^x(x_i, \epsilon_i) = 0 \ \ \forall i \tag{1}$$

where superscripts denote partial derivatives with respect to the superscripted variable.

When there is no regulation, each firm solves

$$\max_{x_i} \quad px_i - c_i(x_i).$$

The first order conditions to the firm's problem show that the profit maximizing choice of output for unregulated firms satisfies

$$x_i^{decent} \ : \ p - c_i^x(x_i) \ \ \forall i. \tag{2}$$

Given that both $c_i(x_i)$ and $f(x_i, \epsilon_i)$ are strictly convex and increasing, we can see that $x_i^{planner} < x_i^{decent}$.

Definition 3. *The* pollution reporting game*, $\mathscr{R}$, is defined by a regulator G; a set of firms $\{i\}_{i=1}^n$; their choices of mechanism, output, and report; the firms' profit functions; and the aggregate state $\mathscr{S}$.*

$$\mathscr{R} = (G, \{i\}_{i=1}^n, \{(T_i, S_i)\}_{i=1}^n, \{(x_i, r_i)\}_{i=1}^n, \pi_i(x_i, r_i; \mathscr{S}, c_i, T_i, S_i))$$

Definition 4. *A set of reports $\{r_i\}_{i=1}^n$ is* truthful *if $r_i = \epsilon_i \ \forall i$.*

Definition 5. *A set of reports from n firms, $\{r_i\}_{i=1}^n$, is* consistent *if it satisfies the adding-up constraint*

$$\sum_{i=1}^n f(x_i, r_i) = D.$$

Without any additional information beyond the aggregate externality stock, the regulator can't distinguish between consistent reports to identify a truthful report, but can be assured that truthful reporting is consistent.

Definition 6. *An* elicited tax *is a mechanism which combines a scoring rule*

$$S_i(f(x_i, r_i), \hat{d}_i, \mathscr{S})$$

and a tax rule

$$T_i(f(x_i, r_i), \hat{d}_i, \mathscr{S})$$

to elicit an agent's private information r_i and levy a tax based on that information, an estimate of that information $\hat{d}_i$, and an observable state $\mathscr{S}$.

In order for an elicited tax to elicit truthful reports, the scoring rule must be proper in the sense described in Brier (1950). Additionally, the tax levied on firm i must not depend on i's own report. If the scoring rule is not proper or firm i's tax depends on i's report, there can be an incentive for i to lie to the mechanism in order to affect its own tax.

Definition 7. *An elicited tax is* strictly proper *if*

1. *the tax penalty is weakly decreasing in the score;*

2. *all reports which minimize agents' scores are consistent;*

3. *all consistent reports minimize agents' scores; and*

4. *the tax penalty for an individual firm with a minimized score contains the firm's true externality generated and a function which does not depend on the firm's report or output choice.*

The definition of strict proper elicited taxes extends the notion of strictly proper scoring rules, and ensures that when the score is minimized firms face the correct marginal incentives to minimize externality generation. As long as the regulator can only verify agents' reports against a function of other reports and an aggregate outcome, there can be multiple score-minimizing sets of reports.

The existence of many consistent lies is related to the notion of elicitation complexity. The complexity of the entire marginal distribution of externality generation may have higher complexity than the mechanism can elicit, given the observable information in the state which can be used to verify the report. Lambert, Pennock, and Shoham in Lambert, Pennock, and Shoham (2008) show that not all properties of a distribution are elicitable, and discuss the complexity of some basic properties of a distribution. The problem we consider is akin to using an observation like the mean to elicit marginal probabilities from forecasters. By using reports from other agents, we are able to construct a more informative outcome with which to score reports from any particular agent.

Definition 8. *The* Brier-Pigou *tax is an elicited tax where the scoring rule is a modified Brier score*

$$S_i(f(x_i, r_i), \hat{d}_i, D) = \left(\frac{f(x_i, r_i) - \hat{d}_i}{D} \right)^2,$$

which uses the regulator's estimate of firm i's marginal externality

$$\hat{d}_i = D - \sum_{j \neq i} f(x_j, r_j),$$

and the tax rule is a score-weighted convex combination of the regulator's estimate and the full damages, i.e.

$$T_i(f(x_i, r_i), \hat{d}_i, D) = (1 - S_i)\hat{d}_i + S_i D$$

The Brier score is the oldest scoring rule in the information elicitation literature Brier (1950). The fact that Brier scores are between zero and one is convenient in constructing the tax rule, since it gives us a smooth weight for a convex combination of the marginal penalty and the punishment.

Definition 9. *Under the Brier-Pigou tax, firms solve*

$$\begin{aligned}
&\max_{x_i, r_i} \; \pi_i(x_i, r_i) \\
&= \max_{x_i, r_i} \; px_i - c_i(x_i) - T_i(f(x_i, r_i), \hat{d}_i, D) \\
&= \max_{x_i, r_i} \; px_i - c_i(x_i) - (1 - S_i(f(x_i, r_i), \hat{d}_i, D))\hat{d}_i - S_i(f(x_i, r_i), \hat{d}_i, D)D
\end{aligned}$$

3 Results

3.1 Analytical results

3.1.1 Strictly proper elicited taxes

Proposition 1. *Under a strictly proper elicited tax, all Nash equilibria of $\mathscr{R}$ are consistent, and all consistent reports are Nash equilibria.*

Proof. 1. All Nash equilibria have consistent reports: Suppose we have a Nash equilibrium. Then $\{(x_i^*, r_i^*)\}_{i=1}^n : \pi_i(x_i^*, r_i^*) \geq \pi_i(x_i, r_i) \;\; \forall \; (x_i, r_i) \neq (x_i^*, r_i^*), \;\; \forall \; i$. By the definition of a strictly proper elicited tax, profits are weakly decreasing in the score, i.e.

$$\frac{\partial \pi_i}{\partial S_i} \leq 0,$$

and the score is minimized only by consistent reports. In any Nash equilibrium, firms' scores must be minimized, or else a firm could weakly improve its profits by changing its reports to make the reports consistent. Therefore, if we have a Nash equilibrium, we must also have consistent reports.

2. All consistent reports are Nash equilibria: Suppose we have a set of consistent reports. By the definition of a strictly proper elicited tax, any deviation from consistency must weakly decrease firms' profits. Since there is no profitable deviation from consistent reporting for any firm, consistent reporting must be a Nash equilibrium. □

Intuitively, if an elicited tax is strictly proper then the only way firms can minimize their scores is by ensuring consistency. Since any inconsistency weakly decreases profit, there is always a profitable deviation towards consistency, and never a profitable deviation away from consistency.

Proposition 2. *If the elicited tax is strictly proper, all Nash equilibria in the reporting game have socially optimal output and aggregate externality production, i.e.,*

$$x_i^{decent} = x_i^{planner}, \text{ and } D_{decent} = D_{planner}.$$

Proof. First, by proposition 1, if a set of output and report choices are a Nash equilibrium, then the reports must be consistent. By the definition of a strictly proper elicited tax, in any Nash equilibrium a firm's first order condition for output must be aligned with the social planner's first order condition for output. Therefore output production is socially efficient. Since there is a one-to-one mapping between output production and externality production, aggregate externality production is also socially efficient.

□

Proposition 2 guarantees that aggregate externality and marginal goods production in any Nash equilibrium will be socially optimal, but it does not guarantee that the marginal taxes will be efficient. In fact, marginal taxes will be inefficient in any non-truthful Nash equilibrium, since the actual tax that a firm pays will include a function which does not depend on that firm's choices. This function will reflect the aggregate lie which the firm must cover for in order for the reports to be consistent.

Given proposition 1 and the definition of a strictly proper elicited tax, from any firm i's perspective $\mathscr{R}$ is effectively a two-player game where i plays against the group of firms who are not i. To address equilibrium selection given the multiplicity of Nash equilibria, we develop a definition of risk-dominant truthful reporting for strictly proper elicited taxes. Theoretical and experimental evidence suggests that risk-dominance is a plausible equilibrium selection mechanism in coordination games with multiple Nash equilibria Harsanyi and Selten (1988).

Definition 10. *Truthful reporting* risk-dominates *consistent lies under a strictly proper elicited tax if*

$$[\pi_i(L,T) - \pi_i(T,T)][\pi_{-i}(L,T) - \pi_{-i}(T,T)] \geq [\pi_i(T,L) - \pi_i(T,T)][\pi_{-i}(T,L) - \pi_{-i}(T,T)],$$

where $-i$ represents the group of players who are not i, $\pi_i(T,T)$ represents player i's payoff when i and $-i$ report truthfully, $\pi_i(L,T)$ represents player i's payoff when i reports a consistent lie and $-i$ report truthfully, and the other terms are defined similarly.

The condition says that the product of losses from each player deviating from truthful reporting is weakly greater than the product of losses from each player being deviated against from truthful reporting. Intuitively, it should be weakly worse to be the liar than to be the one who is lied about.

Note that this definition of risk-dominance does not imply that truthful reporting would be risk-dominant for all players who are not i, but rather that truthful reporting is risk-dominant for any player i and that *aggregate* truth telling is risk-dominant for the group of players who are not i. Within the group of $-i$ players, there may be players for whom truthful reporting against the remaining players is not risk-dominant.

3.1.2 The Brier-Pigou tax

Lemma 1. *Under the Brier-Pigou tax, any set of consistent reports will result in all firms receiving a score of zero. Conversely, if the reports are inconsistent, all firms will receive positive scores. Further, all firms will have the same scores under any set of reports.*

Proof. 1. Any set of consistent reports results in a zero score for all firms: Suppose a set of consistent reports, $\{r_i^*\}_{i=1}^n : \sum_{i=1}^n f(x_i, r_i^*) = D$. Then $\sum_{i=1}^n f(x_i, r_i^*) = f(x_i + r_i^*) + \sum_{j \neq i} f(x_j, r_j^*) = D,\ \forall i, j$. Firm i's score under these reports is

$$\begin{aligned} S_i &= \left(\frac{f(x_i, r_i^*) - \hat{d}_i}{D} \right)^2 \\ &= \left(\frac{f(x_i, r_i^*) - D + \sum_{j \neq i} f(x_j, r_j^*)}{D} \right)^2 = 0, \end{aligned}$$

which is true for all i. This proves the first part of the lemma.

2. If the reports are inconsistent, all firms will receive positive scores: Suppose a set of inconsistent reports, $\{\tilde{r}_i\}_{i=1}^n : \sum_{i=1}^n f(x_i, \tilde{r}_i) \neq D$. Firm i's score under these reports is

$$
\begin{aligned}
S_i &= \left(\frac{f(x_i, \tilde{r}_i) - \hat{d}_i}{D}\right)^2 \\
&= \left(\frac{f(x_i, r_i^*) - D + \sum_{j \neq i} f(x_j, \tilde{r}_j)}{D}\right)^2 > 0,
\end{aligned}
$$

which is true for all i. This proves the second part of the lemma.

3. All firms have the same scores under any set of reports:

$$
\begin{aligned}
S_i &= \left(\frac{f(x_i, r_i) - \hat{d}_i}{D}\right)^2 \\
&= \left(\frac{f(x_i, r_i) - D + \sum_{j \neq i} f(x_j, r_j)}{D}\right)^2 \\
&= \left(\frac{\sum_{i=1}^n f(x_i, r_i) - D}{D}\right)^2 && = S_j,
\end{aligned}
$$

which is true for all i. This proves the third part of the lemma. □

The symmetry of firms' scores comes from the symmetry of the Brier score, and the fact that the regulator only has a single common piece of information with which to score firms.

Lemma 2. *Under perfect competition in output and no collusion, the Brier-Pigou tax is a strictly proper elicited tax.*

Proof. From the definition of the Brier-Pigou tax, the tax penalty is weakly decreasing in the score. By lemma 1, the Brier-Pigou tax satisfies conditions two and three for strict propriety. All that remains is to check the fourth condition.

By lemma 1, if the reports are consistent then every firm must have a score of zero, and from the structure of the Brier-Pigou tax the minimum score is zero. Then, conditional on choosing a consistent report, each firm solves

$$
\max_{x_i} p x_i - c_i(x_i) - \hat{d}_i,
$$

where

$$
\begin{aligned}
\hat{d}_i &= D - \sum_{j \neq i} f(x_j, r_j) \\
&= \sum_{i=1}^{n} f(x_i, \epsilon_i) - \sum_{j \neq i} f(x_j, r_j) \\
&= f(x_i, \epsilon_i) + \sum_{j \neq i} [f(x_j, \epsilon_j) - f(x_j, r_j)] \\
&= f(x_i, \epsilon_i) + \ell_i,
\end{aligned}
$$

where ℓ_i is the aggregate lie which firm i must account for in its report. Since the firms are perfectly competitive, ℓ_i doesn't depend on x_i or r_i. So, under perfect competition in output and no collusion, the Brier-Pigou tax is a strictly proper elicited tax. □

Proposition 3. *Under the Brier-Pigou tax, the perfectly collusive profit maximizing choices of output and reports is not consistent, not a Nash equilibrium, and not socially optimal.*

Proof. A perfectly collusive group of firms facing a world price for their output solves

$$
\begin{aligned}
\max_{\{x_i\},\{r_i\}} \quad & \sum_{i=1}^{n} (p * x_i - c_i(x_i) - T_i(D, S(D, \{x_i\}, \{r_i\}), f(x_i, \hat{\epsilon}_i))) \\
\text{s.t. } & T_i = (1 - S_i)\hat{d}_i + S_i * D, \\
& S_i = \left(\frac{f(x_i, r_i) - \hat{d}_i}{D} \right)^2, \\
& \hat{d}_i = D - \sum_{j \neq i} f(x_j, r_j)
\end{aligned}
$$

Taking derivatives, the first order conditions are

$$
\begin{aligned}
x_i :& p - c_i^x(x_i) - (1 - S_i) f_x(x_i, \hat{\epsilon}_i) + \frac{\partial S_i}{\partial x_i}(D - f(x_i, \hat{\epsilon}_i)) \\
&+ \sum_{j \neq i} \left[-f_x(x_j, \epsilon_j)(1 - 2S_j) + f_x(x_j, r_j)(1 - S_j) + \frac{\partial S_j}{\partial r_i}(D - f(x_j, \epsilon_j)) \right] = 0 \\
r_i :& - \frac{\partial S_i}{\partial r_i}(D - f(x_i, \hat{\epsilon}_i)) \\
&- \sum_{j \neq i} \left[\frac{\partial f(x_j, \hat{\epsilon}_j)}{\partial r_i}(1 - S_j) + \frac{\partial S_j}{\partial r_i}(D - f(x_j, \hat{\epsilon}_j)) \right] = 0
\end{aligned}
$$

Inspection reveals that consistent reporting will not solve these equations. By proposition 1, the collusive solution is not a Nash equilibrium of the underlying noncooperative game between the firms. By proposition 2, the collusive solution is therefore socially inefficient as well.

□

The intuition for this result is related to the definition of a strictly proper elicited tax. If an elicited tax is strictly proper for non-collusive reporting firms, firm i's tax penalty can depend on $-i$'s reports and outputs. When the firms are not colluding, they won't internalize their effects on others and will be incentivized to produce consistent reports. When the firms are colluding, their incentive to be consistent is weakened because they will internalize the effects of their reports on others. Simulation results in the next section show that the welfare loss due to collusive reporting is likely to be bounded and small.

3.2 Simulation results

In this section, we present simulations of the Brier-Pigou tax with two firms using following functional form assumptions on costs and damages:

$$c_i(x_i) = a_{i1}x_i + a_{i2}x_i^2$$
$$f(x_i, \epsilon_i) = (x_i\epsilon_i)^2.$$

The price of the output good is normalized to 1 in all cases.

3.2.1 Welfare

The primary benefit of the Brier-Pigou tax is that it improves social welfare over the unregulated market outcome. In this section, we simulate social welfare under different marginal cost parameter realizations and reporting behaviors to illustrate the mechanism's performance and study the effects of reporting collusion.

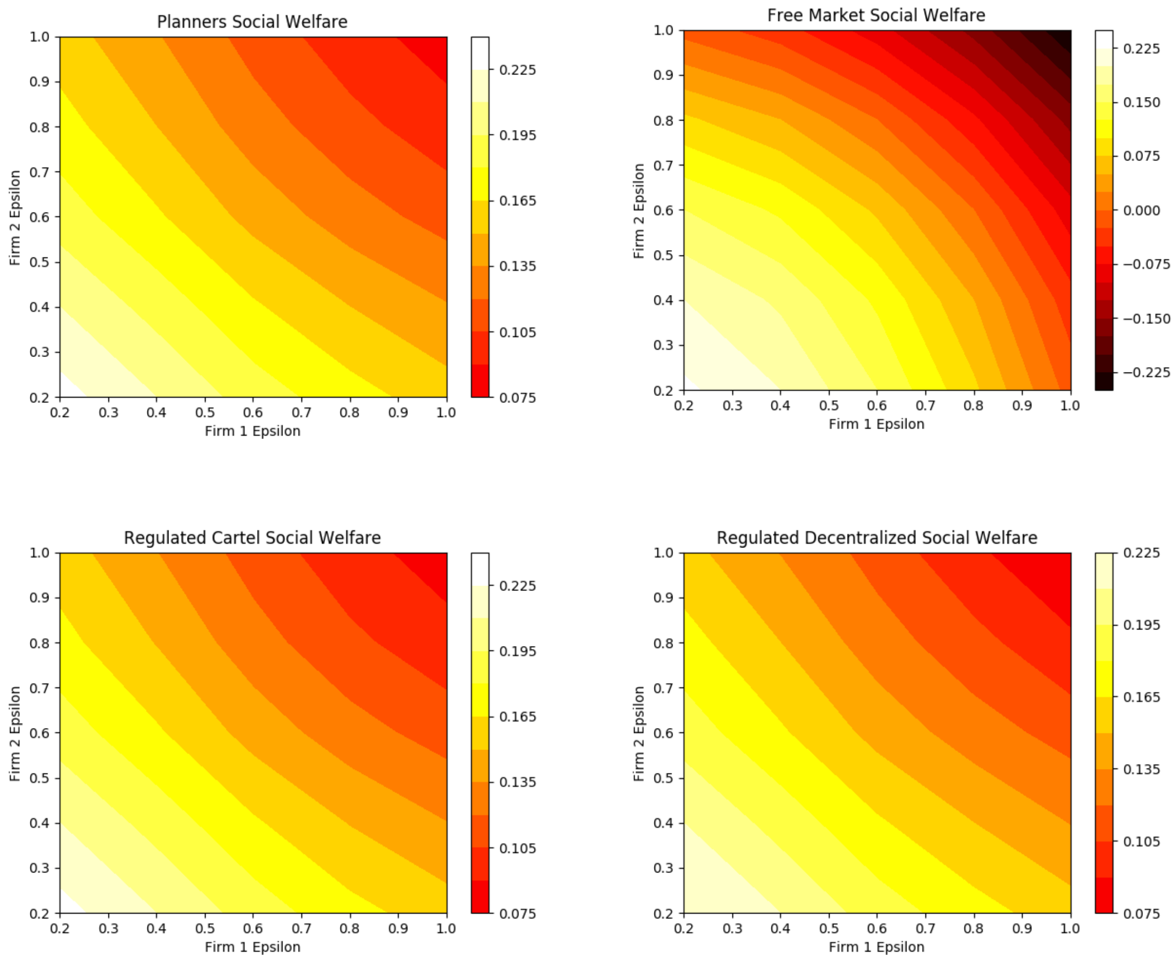


Figure 1: $a_{11} = 0.5, a_{12} = 0.5, a_{21} = 0.5, a_{22} = 0.5$

Figure 1 shows welfare as heat against ϵ_1 and ϵ_2, with hotter colors representing higher welfare. The planner's heatmap shows the underlying welfare possibility frontier (value of equation 2 under solutions to equation 1), and the free market heatmap shows the floor of the welfare possibility set (value of equation 2 under solutions to equation 2).

The level sets of social welfare given socially optimal production are concave, while the level sets of social welfare given unregulated profit maximizing production are convex. This convexity inversion highlights the welfare effects of internalizing D in solving equations 1 and 2, respectively. In solving equation 2, firms do not account for D at all. Under an elicited tax, the effects of marginal production decisions on social welfare enter the firm's problem through the tax penalty. When the elicited tax is strictly proper, the level sets of the firms' objective functions are aligned with the planner's. Even under perfect collusion, the Brier-Pigou tax is able to bring profit maximizing output decisions significantly closer to the social optimum than the unregulated market is able to achieve.

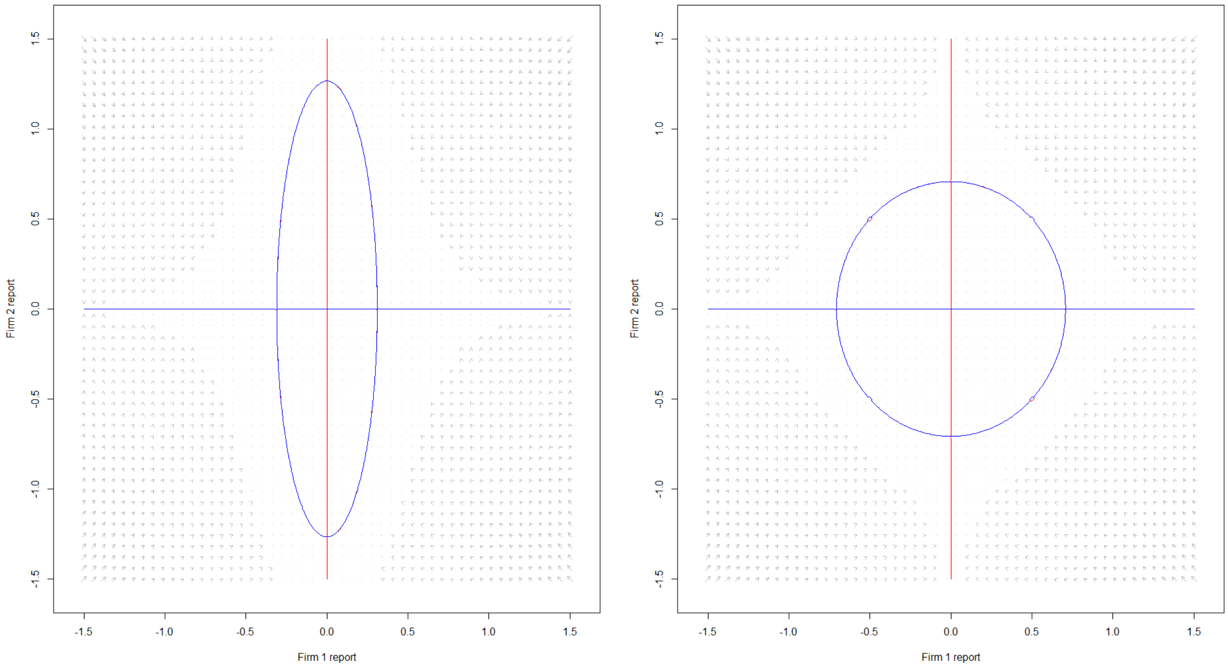
Figure 2 shows the average social welfare across all realizations of ϵ_1 and ϵ_2 for different marginal cost parameters. The planner's and free market welfares are normalized to 1 and 0 respectively in each row.

E[Social Welfare] vs. Marginal Firm Costs	Planner	Free	Decentralized	Cartel
A11: 0.25, A12: 0.25, A21: 0.75, A22: 0.75	1	0	0.981	0.993
A11: 0.75, A12: 0.25, A21: 0.25, A22: 0.75	1	0	0.981	0.990
A11: 0.25, A12: 0.75, A21: 0.75, A22: 0.25	1	0	0.977	0.987
A11: 0.5, A12: 0.5, A21: 0.5, A22: 0.5	1	0	0.992	0.996

Figure 2

Figure 2 shows that the Brier-Pigou tax significantly improves social welfare over the free market outcome across a wide range of marginal cost structures between firms.

While proposition 2 guarantees that Nash equilibria will be socially optimal, not all initial values lead the solver to Nash equilibria. As the values of ϵ_i change, the set of paths leading the solver from any given initial value to a Nash equilibrium also change. In particular, initial values that are large lead to non-Nash corner solutions in production, and initial values close to 0 may not converge. Figure 3 illustrates this problem with vector fields for the firms' reporting best response functions under fixed levels of output and common marginal costs.



(a) (b)

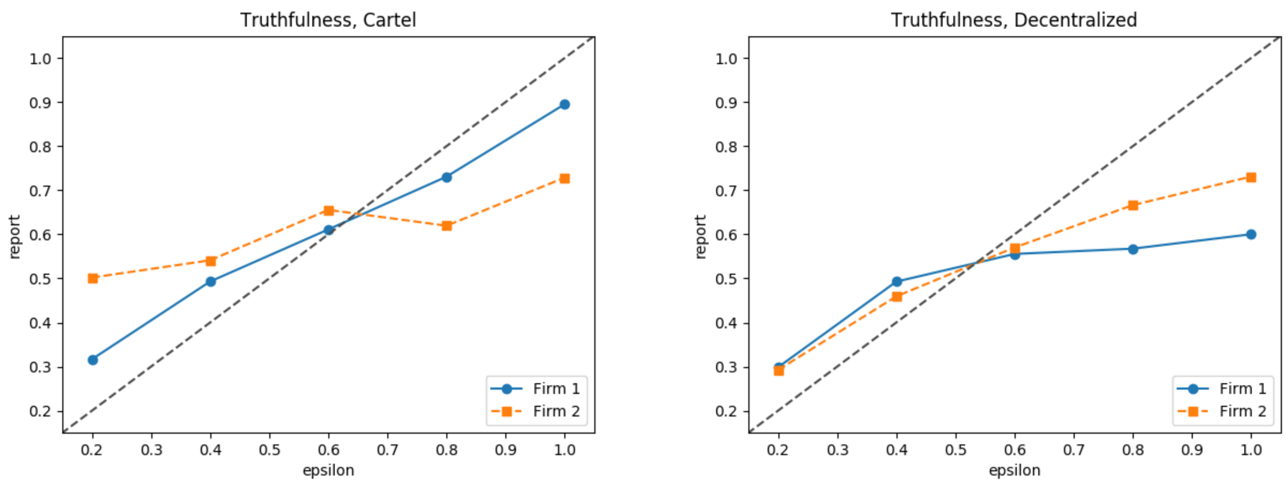
Figure 3: Vector field of firm 1 (red) and firm 2 (blue) reporting best-response functions when (a) $\epsilon_1 = 0.25, \epsilon_2 = 0.75$ and (b) $\epsilon_1 = 0.5, \epsilon_2 = 0.5$. Both functions intersect on the blue circle, giving the set of consistent reports given D.

To correct for this, we calculate welfare at each point in figure 1 by averaging over a grid of initial values for reporting. The non-convergent initial values cause some distortion in the welfare values for a particular point, but the mechanism's performance is still evident. Numerical stability can be assured with bounds on the initial values informed by lemma 3, and the social welfare under the mechanism and no collusion should match the planner's exactly. Proposition 3, however, suggests that the gap between social welfare under the planner and social welfare under perfect collusion is not a numerical artifact, but rather the result of strategic reporting behavior by firms against the regulator.

3.2.2 Truthfulness

Proposition 2 guarantees that Nash equilibria of a strictly proper elicited tax will be socially efficient and will elicit consistent reports, but there is no guarantee that it will elicit truthful reports. The multiplicity of Nash equilibria with consistent lying suggests the distribution of profits among firms may not be socially optimal, which could result distorted long-run entry and exit decisions. While we do not model entry and exit here, we offer simulation evidence about the nature of equilibrium selection under the Brier-Pigou tax.

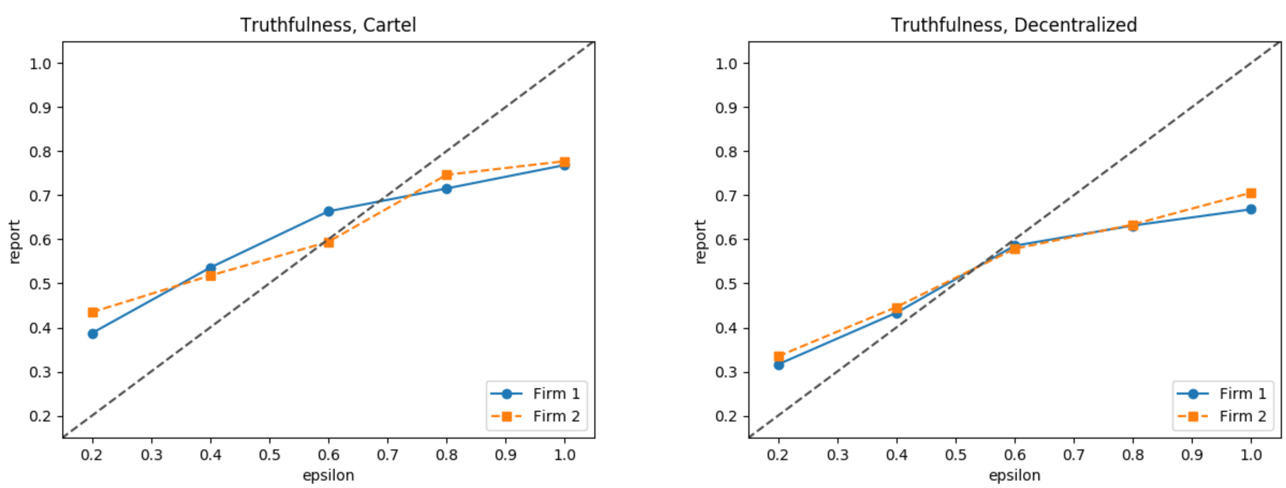
Figures 4 and 5 show the deviations in truthful reporting for each firm given their ϵ_i under perfect collusion and perfect competition. There appears to be a systematic pattern in reporting as a function of the marginal externality each firm is responsible for. Firms creating smaller shares of the externality tend to overreport their contributions, while firms creating larger shares of the externality tend to underreport their contributions.



(a) Under perfect reporting collusion

(b) Under no reporting collusion

Figure 4: $a_{11} = 0.5, a_{12} = 0.1, a_{21} = 0.1, a_{22} = 0.2$



(a) Under perfect reporting collusion

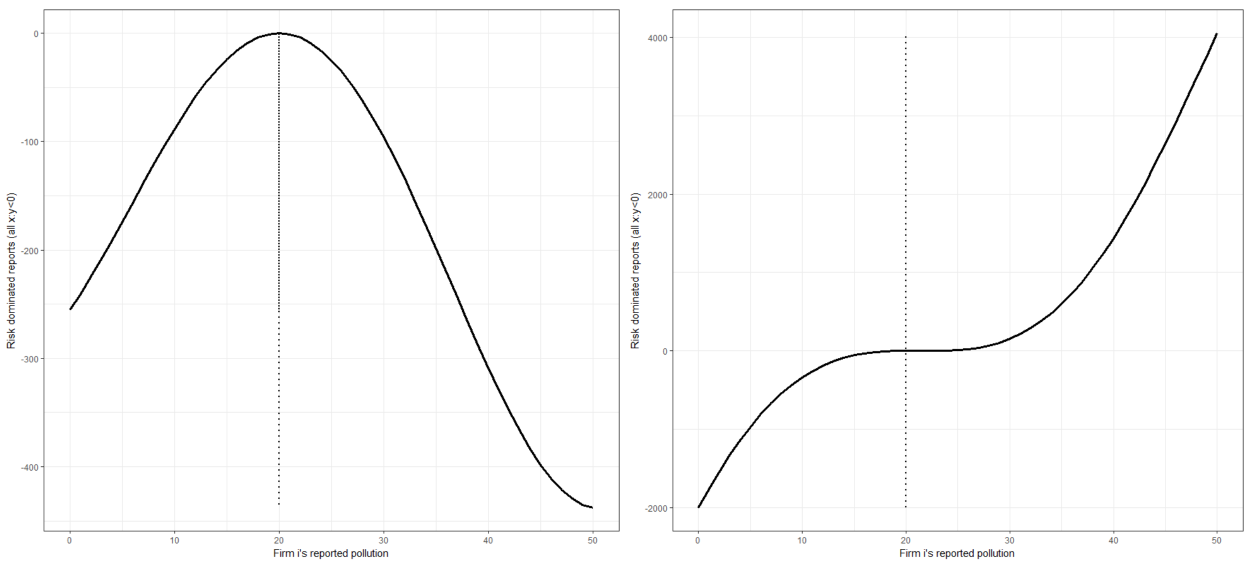
(b) Under no reporting collusion

Figure 5: $a_{11} = 0.5, a_{12} = 0.5, a_{21} = 0.5, a_{22} = 0.5$

Risk dominance offers one rationalization for this behavior. Truthful reporting to the Brier-Pigou tax will risk-dominate consistent lies which satisfy the inequality

$$
\begin{aligned}
(2\hat{d}_i^T - D)\left\{\ell\left(1+\left(\frac{\ell}{D}\right)^4\right)+\ell+\frac{\ell^3}{D^2}\right\} - \ell^2\left(1+\left(\frac{\ell}{D}\right)^4\right) \\
+2\left(\frac{\ell^4}{D^2}\right)+\left(\frac{\ell^3}{D^2}\right)(n-2)D\left(1+\left(\frac{\ell}{D}\right)^2\right) \leq 0 \qquad (3)
\end{aligned}
$$

where $\hat{d}_i^T$ is firm i's true marginal externality generation, $\hat{d}_i^L$ is firm i's reported marginal externality generation under a consistent lie, and $\ell = \hat{d}_i^L - \hat{d}_i^T$ is the size of the lie. Inspecting this inequality suggests that the risk-dominance of truthful reporting will depend on i's true share of the aggregate measured externality ($\hat{d}_i^T/D$) and the number of firms (n).



(a) $\frac{\hat{d}_i^T}{D} = 0.2, n = 2$ (b) $\frac{\hat{d}_i^T}{D} = 0.2, n = 15$

Figure 6: Risk dominance of truthful reporting under different market sizes. Negative values indicate reports which are risk-dominated by the truth.

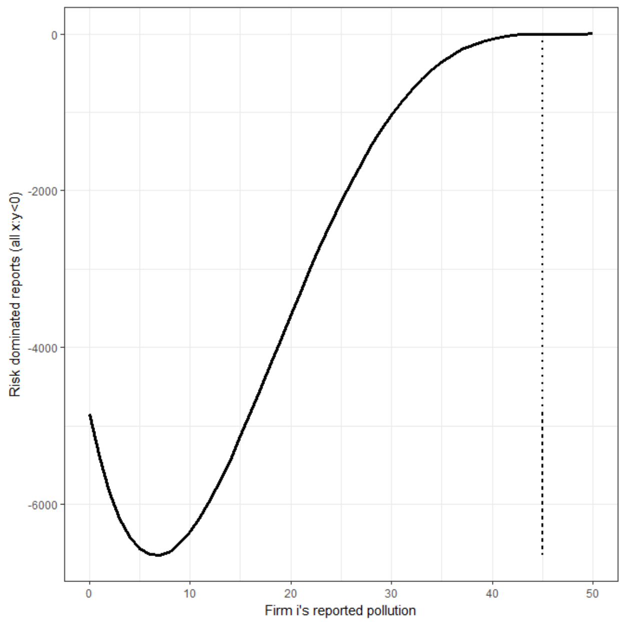
Figures 6 and 7 plot the left hand side of inequality 3 as a function of possible consistent reports ($\hat{d}_i^L$). The dotted vertical line displays firm i's true marginal externality generation ($\ell = 0$). Negative values indicate consistent reports which are risk-dominated by truthful reporting.

Truthful reporting risk-dominates all consistent lies for firms which produce a "large enough" proportion of the aggregate externality (i.e. large enough $\hat{d}_i^T/D$). The threshold for "large enough" appears to be increasing in the size of the market, so that truthful reporting is globally risk-dominant for a firm which produces 20% of D in a $2-$firm market, but not for a firm which produces 20% of D in a $15-$firm market (figure 6). Holding market size constant, however, share of externality generation is a key factor in being "large enough" that truthful reporting is globally risk-dominant (panel 7).

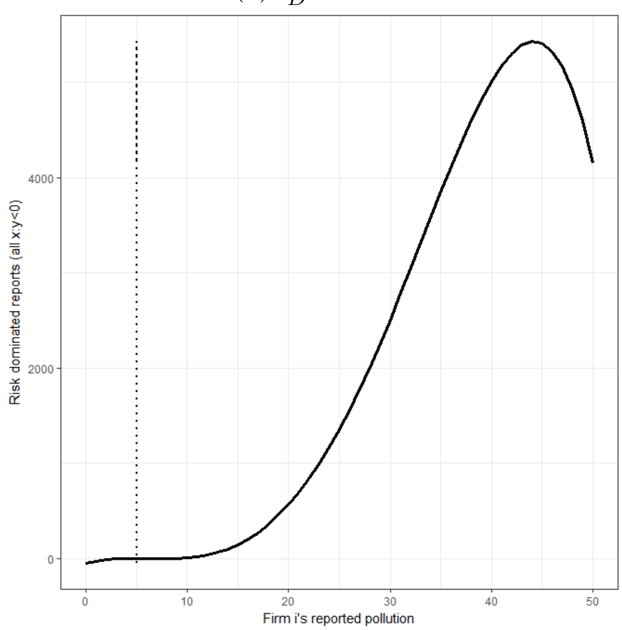
While it may appear that small (i.e. low externality share) firms have no incentive to report truthfully, numerical inspection of inequality 3 reveals that truthful reporting locally risk-dominates consistent lies in a neighborhood of the truth even for very small firms. Additonally, each panel only represents i's incentive to report truthfully. The fact that a small firm has a small incentive to report truthfully does not suggest that it will be able to tell its most-preferred lie - larger firms may face a much stronger incentive to tell the truth, which would force smaller firms to stay near their truths to maintain consistency. Interestingly, consistent underreporting appears to be uniformly risk-dominated by truthful reporting. Since consistency requires that any underreport by one firm must be balanced by an equal overreport by others, if underreporting is uniformly risk-dominated by truthful reporting, large overreports seem less likely.

To explore the relationship between risk-dominant truthful reporting, firm's externality share, and market size, we simulated inequality 3 under different externality shares and market sizes and calculated the proportions of non-truthful reports which did and didn't satisfy inequality 3. This gave us the probability that a randomly selected consistent lie would be risk-dominated by the truth, and the probability that a randomly selected consistent lie would risk-dominate the truth.

For small market sizes ($n \leq 3$), truthful reporting globally risk-dominates consistent lying for all values of $\hat{d}_i^T/D$. As n increases past 3, however, the probability that a randomly selected consistent lie risk-dominates truthful reporting increases for firms of all $\hat{d}_i^T/D$, beginning with firms that have small $\hat{d}_i^T/D$. By $n = 100$, a threshold appears to stabilize at $\hat{d}_i^T/D = 0.5$ where, for $\hat{d}_i^T/D > 0.5$, truthful reporting risk-dominates more than half of all consistent lies. For firms with $\hat{d}_i^T/D = 1$, truthful reporting appears to globally risk-dominate consistent lies no matter the market size.

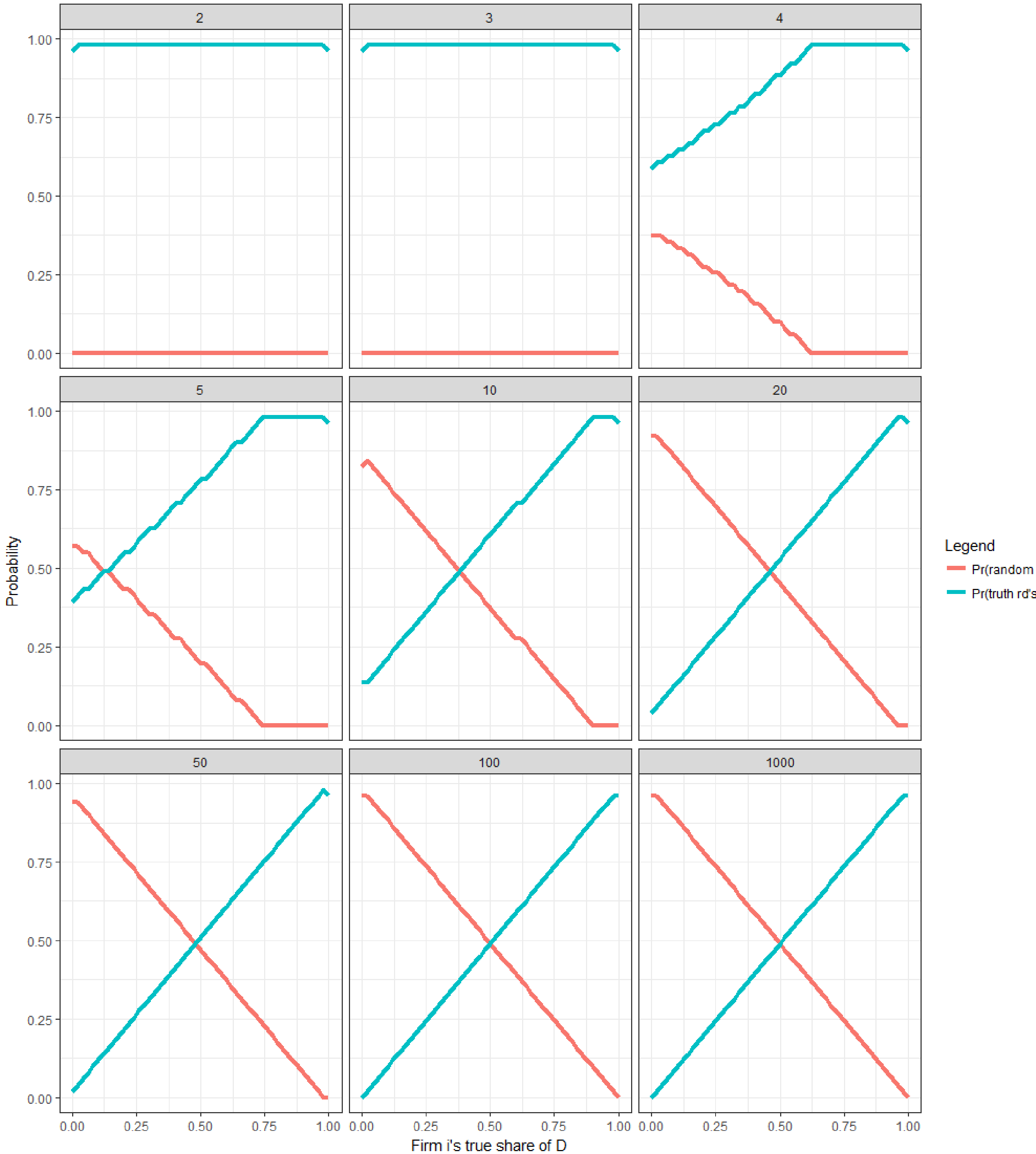


(a) $\frac{\hat{d}_i^T}{D} = 0.9$



(b) $\frac{\hat{d}_i^T}{D} = 0.08$

Figure 7: Risk dominance of truthful reporting with 15 firms. Negative values indicate reports which are risk-dominated by the truth.



2
3
4
5
10
20
50
100
1000
1.00
0.75
0.50
0.25
0.00
Probability
0.00
0.25
0.50
0.75
1.00
Firm i's true share of D
Legend
Pr(random
Pr(truth rd's

Figure 8

Combined with the simulation evidence that perfect collusion will not significantly degrade social welfare and analytical guarantees that consistency will be near the short-run social optimum for perfectly competitive output markets, the results in figure 8 suggest that the Brier-Pigou tax may be well suited to regulating small groups of firms at a time.

4 Future research

While scoring rules and corrective taxes are well-understood on their own, little is known about the properties of elicited taxes. The Brier-Pigou tax is an intuitive mechanism, combining a well-understood scoring rule with a simple smooth tax penalty, but the space of elicited taxes is large. From our perspective, the most important questions about elicited taxes are: (1) Are there are elicited taxes which result in optimal marginal taxes across firms in addition to aggregate welfare optimality? (2) Would such mechanisms allow the regulator to learn the firms' marginal contributions to the externality? (3) How would such mechanisms perform under collusion in reporting? We suggest two modifications to the Brier-Pigou tax which may help answer these questions.

Scoring rules: The symmetry of the Brier scoring rule may prevent truthful reporting from being risk-dominant. Asymmetric scoring rules, such as the log score, may be able to make truthful reporting more attractive.

Tax rules: The Brier-Pigou tax uses a score-weighted convex combination of estimated marginal damages and total damages to induce consistent reports. Allowing for more flexible interactions between the score, the aggregate information, and the regulator's estimate of marginal damages may help induce truthfulness.

5 Conclusion

In this paper, we propose a promising new class of mechanism for solving externality problems under asymmetric information. Elicited taxes use techniques from information elicitation to assess corrective taxes to solve problems like nonpoint pollution control, and provide a method by which regulators can apply fair penalties without full information regarding marginal damages. We show that any elicited tax which is strictly proper satisfies certain desirable qualities, and explore the Brier-Pigou tax in detail through simulations. We hypothesize that there exist some scoring rules which may lead to truthful consistent reporting as a unique or highly attractive Nash equilibrium.

Analytical results show that the Nash equilibria of the Brier-Pigou tax significantly improve social welfare over deregulated outcomes. Numerical experiments show that the tax significantly improves social welfare on average even when non-equilibrium outcomes are considered. Under perfect reporting collusion, the mechanism results in welfare that is very competitive with respect to the planner's outcome. Across all of our simulations, the expected welfare under the mechanism strictly dominates the expected welfare under the free market.

When the reports are consistent, the tax levied under a strictly proper elicited tax actually achieves the social planner's welfare in the Nash equilibrium under no collusion. We show analytically that in the presence of reporting collusion, the reports will not be consistent. Simulations show that the resulting welfare loss is extremely small. Such collusion results in excess pollution, which is balanced out by additional production which is still socially beneficial. Even more encouraging is that this overproduction tends to be allocated so as to reduce its negative effects on welfare. Due to the inconsistency of collusive reports, the regulator will know that the firms are lying. Future work may develop a method to use this feature of strictly proper elicited taxes to prevent collusion in the first place.

Simulations show that even when firms report consistent lies to the Brier-Pigou tax, the lies are usually not egregious - deviations from the truth are relatively small. Numerical analysis suggests that truthfulness risk-dominates small lies and underreporting, and that firms producing larger shares of the externality have stronger incentives to be report truthfully. The elicited tax also tends to be budget-balanced, meaning that the taxes levied cover the damages caused. While it is exactly budget-balanced in any Nash equilibrium, our welfare simulations suggest that it is close to budget-balanced out of Nash equilibria as well.

Standard pollution control devices such as direct regulation of emissions or production are not appropriate for problems characterized by asymmetric information regarding marginal damages. When the regulator doesn't know each firm's marginal contributions to the aggregate externality, they are unable to tax them optimally. Existing mechanisms may not be budget-balanced, or may not use observable information and individual reports fully. Elicited taxes offer a promising way to combine elicitation with corrective taxation to solve problems in environmental economics characterized by such asymmetric information. These mechanisms have several advantages, including robustness to collusion, a variety of ways to incorporate prior information, and the ability to reveal truthful or near-truthful information to the regulator.

References

Baerenklau, K. A. (2002): "Green Payment Programs for Nonpoint Source Pollution Control: How Important is Targeting for Cost-Effectiveness?," *Journal of Agricultural and Resource Economics*, 27, 406–419.

Brier, G. W. (1950): "VERIFICATION OF FORECASTS EXPRESSED IN TERMS OF PROBABILITY," *Monthly Weather Review*, 78(1), 1–3.

Groves, T. (1973): "Incentives in Teams," *Econometrica*, 41, 617–631.

Hansen, L. G. (1973): "A Damage Based Tax Mechanism for Regulation of Non-Point Emissions," *Environmental and Resource Economics*, 12, 99–112.

Harsanyi, J. C., and R. Selten (1988): *A General Theory of Equilibrium Selection in Games*. MIT Press.

Holmstrom, B. (1982): "Moral Hazard in Teams," *Econometrica*, 13, 324–340.

Kritikos, A. S. (2004): "A penalty system to enforce policy measures under incomplete information," *International Review of Law and Economics*, 24, 385–403.

Lambert, N. S., D. M. Pennock, and Y. Shoham (2008): "Eliciting Properties of Probability Distributions," in *Proceedings of the 9th ACM Conference on Electronic Commerce*, EC '08, pp. 129–138, New York, NY, USA. ACM.

Segerson, K. (1988): "Uncertainty and Incentives for Nonpoint Pollution Control," *Journal of Environmental Economics and Management*, 15, 87–98.

6 Appendix

Lemma 3. *Nash equilibria of $\mathscr{R}$ under the Brier-Pigou elicited tax exist.*

Proof. From definition 9, a firm's first order conditions for output and reporting are

$$\begin{aligned} x_i &: p - c_x(x_i) - (1 - S_i)\hat{d}_i^x + S_i^x(D - \hat{d}_i) \leq 0 \\ r_i &: \ - S_i^r(D - \hat{d}_i) \leq 0 \end{aligned}$$

where superscripts indicate partial derivatives with respect to the superscripted variables.

Let $\bar{x}$ be defined as

$$\max\{x \geq 0 : px_i - c(x_i) - (1 - S_i)\hat{d}_i - S_i D\},$$

and $\bar{r}$ be defined as

$$r : f(\underline{x}_i, r) = D,$$

where $\underline{x}_i$ is the smallest amount of output produced by a single firm in the market. Both bar $\bar{x}$ and $\bar{r}$ are finite upper bounds on production and the reports. Since x_i and r_i are nonnegative, we can restrict them to lie in the compact and convex intervals $[0, \bar{x}]$ and $[0, \bar{r}]$, respectively. From definition 8, the first order conditions above describe continuous functions. Applying the Brouwer Fixed Point Theorem, at least one Nash equilibrium of the $2n$ equations described above exists. $\square$