6-2019

Adversarial contract design for private data commercialization

Parinaz NAGHIZADEH

Arunesh SINHA
Singapore Management University, aruneshs@smu.edu.sg

Citation

NAGHIZADEH, Parinaz and SINHA, Arunesh. Adversarial contract design for private data commercialization. (2019). *Proceedings of the 20th ACM Conference on Economics and Computation (EC-19), Phoenix, Arizona, USA, 2019 June 24–28*. 681-699.
Available at: https://ink.library.smu.edu.sg/sis_research/4797

Adversarial Contract Design for Private Data Commercialization

PARINAZ NAGHIZADEH*, Purdue University
ARUNESH SINHA*, University of Michigan

The proliferation of data collection and machine learning techniques has created an opportunity for commercialization of private data by data aggregators. In this paper, we study this data monetization problem as a mechanism design problem, specifically using a contract-theoretic approach. Our proposed adversarial contract design framework provides a fundamental extension to the classic contract theory set-up in order to account for the heterogeneity in honest buyers' demands for data, as well as the presence of adversarial buyers who may purchase data to compromise its privacy. We propose the notion of *Price of Adversary* (*PoAdv*) to quantify the effects of adversarial users on the data seller's revenue, and provide bounds on the *PoAdv* for various classes of adversary utility. We also provide a fast approximate technique to compute contracts in the presence of adversaries.

Additional Key Words and Phrases: contract theory; pricing private data; data commercialization

ACM Reference Format:
Parinaz Naghizadeh and Arunesh Sinha. 2019. Adversarial Contract Design for Private Data Commercialization. In *The 20th ACM conference on Economics and Computation (EC '19), June 24–28, 2019, Phoenix, AZ, USA.* ACM, New York, NY, USA, 19 pages. https://doi.org/110.1145/3328526.3329633

1 INTRODUCTION

The large-scale adoption of data-driven decision making by businesses has led to a boom in big data collection and analysis techniques. With increasing amount and demand for data, companies have found a business opportunity in offering data-based services to other companies, or selling their data to interested parties [30, 32]. Interest in data monetization is evidenced by the rise of *data marketplaces*, where firms and individuals can buy, sell, or trade, second or third party data. Examples include Salesforce's Data Studio, Oracle's BlueKai, and Adobe's Audience Marketplace. Data commercialization faces many challenges, including IP protection, liability, pricing, and preserving privacy [32]. In this paper, we focus on the latter two challenges of pricing and privacy.

The challenge of pricing refers to the fact that to accommodate diverse demands, data sellers offer different plans and pricing to their buyers. Even with identical data, buyers may derive different benefits from utilizing it, e.g., due to different expertise, or how this data complements the buyer's existing knowledge. Therefore, to maximize revenue, the data seller should account for this demand diversity by packaging its data accordingly. Further, despite its revenue benefits, data commercialization has to overcome the challenge of limiting privacy risks for the data subjects in the database. Specifically, adversarial buyers can request access to the database, attempting to

*Both authors contributed equally to this research.

EC '19, June 24–28, 2019, Phoenix, AZ, USA

ACM ISBN 978-1-4503-6792-9/19/06…$15.00
https://doi.org/110.1145/3328526.3329633

compromise the privacy of the data subjects. Therefore, data sellers should account for this risk when designing and pricing data plans.

Mechanism design is a natural approach for the design problem stated above. In this paper, we take a contract theoretical approach [20], a well-known research area within mechanism design[1], to address both the aforementioned pricing and privacy challenges of data commercialization by proposing the design of a set of contracts with varying privacy levels. In general, contracts are designed in the presence of two types of informational asymmetry between the principal and the agents: hidden action (unobservable actions of agents) and hidden information (unobservable types of agents). We work in the hidden information setting here. Contract theory with hidden information, in the classic context of pricing of goods [21], is the study of principal-agent problems, in which the principal (here, the data seller) designs a set of contracts with varying consumption level so as to extract maximum revenue from agents (buyers) with unknown types. We build on this framework, and further extend it by introducing a mixture of honest and adversarial buyers.

We study the problem of pricing a bundle of queries at different privacy levels with the aim of (a) maximizing revenue by offering different prices for varying privacy levels in order to accommodate the diversity of demands for the query bundle, and (b) accounting for the risks from adversarial users by modifying the contracts' pricing accordingly. We use the well accepted *ϵ-differential privacy* concept as the measure of privacy [9]. We make an effort to keep our design practical by attempting to adhere to practices already in place in data marketplaces (see Sections 2 and 3).

Technical contributions: Our contributions can be summarized as follows:

(1) We provide a fundamental extension of the classical contract theory with hidden information framework by introducing buyer types with misaligned incentives (honest and adversarial buyer types). We cast the private data selling problem in this new contract theory framework. We analyze the properties of this new kind of contract design problem.
(2) We compare the structural properties of the optimal contracts when the adversarial buyers are present with those of the optimal contracts when all buyers are honest. In particular, existing contract theory results suggest that given $n + 1$ types of agents (n types of honest buyers and an adversarial type), the principal should design up to $n + 1$ contracts. We nonetheless show that the data seller will offer at most n contracts (Lemma 3.1) in the adversarial setting. In other words, it is optimal for the data seller to avoid the impractical option of designing a contract for the adversary. We further show that despite this difference, the optimal contracts for the adversarial settings will continue to satisfy some structural conditions that are equivalent to those of the classic, non-adversarial setting (Theorem 4.2).
(3) We further extend the contract-theory framework by incorporating post-hoc fines (to be collected in case of a privacy breach) in the pricing of query bundles, and analyze their effect on the contract design problem, showing that fines can be helpful in reducing loss due to the adversarial buyers in many situations.
(4) We propose the notion of *Price of Adversary* (*PoAdv*) to quantify the loss incurred by the data owner due to the presence of adversarial data buyers. We show that while *PoAdv* can be unbounded in the worst case (Lemma 4.5), it is possible to bound the *PoAdv* for a large class of problems (Theorem 5.4).
(5) We provide a fast approximate technique to compute the contracts in presence of adversaries (Algorithm 1).

[1] In micro-economics, contract design is studied as a mechanism design problem, however, that is typically not the case in computer science. Mathematically, contracts design has the same optimization nature as other mechanism design problems.

The remainder of this paper is organized as follows. We present background information on data marketplaces and differential privacy in §2. §3 introduces the buyers' models and the data seller's contract design problem. We study the adversarial contract design problem in §4, and present a fast approximation algorithm for solving this problem in §5. We present numerical simulations in §6 and review related work and summarize our work in §7. All omitted and full version of proofs are available in the appendix after the references.

2 BACKGROUND

Database marketing examples: Currently, the two industries leading database marketing are data brokers (who mine and sell consumer data to businesses), and data marketplaces (which provide a platform for buying, selling, and trading data). We elaborate upon typical privacy guarantees offered by each with an example. Among data brokers, Acxiom, one of the largest brokers worldwide, states that they maintain "privacy compliant data" through data encryption and secure data management techniques [1]. On the other hand, Among data marketplaces, the user service agreement of Salesforce Data Studio [29], provides more detailed information about their market structure. For instance, Salesforce states that they use "unique user identifiers (user IDs) to help ensure that activities can be attributed to the responsible individual", and that security logs are kept "in order to enable security reviews and analysis." Our model in Section 3 takes the availability of these monitoring techniques into account. It is clear that following such safe practices is imperative when dealing with private information, e.g., as evidenced by the recent Cambridge Analytica case [14].

Differential privacy: A popular formalism of privacy loss due to queries from statistical databases is that of *differential privacy (DP)* [9, 10]. Formally, let $\mathcal{K}$ be a randomized algorithm used to answer queries from a database, and consider two databases $\mathcal{D}_1$ and $\mathcal{D}_2$ that differ in exactly one entry (row). Then, $\mathcal{K}$ is ϵ-DP for $\epsilon \geq 0$ if for any possible set of output O,

$$Pr(\mathcal{K}(\mathcal{D}_1) \in O) \leq \exp(\epsilon) \cdot Pr(\mathcal{K}(\mathcal{D}_2) \in O) . \tag{1}$$

In words, ϵ-DP requires that the output of $\mathcal{K}$ remains sufficiently unaffected (as quantified by ϵ), whether or not a single data subject's data is included in the database. For continuous-valued queries, a method for achieving differential privacy is the introduction of carefully selected random noise in the responses. Specifically, let f be a query function, returning the true value $f(D)$ on database D. In order to guarantee ϵ-DP, an algorithm K can introduce additive Laplacian noise, returning instead $f(D) + \text{Lap}(\Delta f/\epsilon)$, where Δf is the sensitivity of the query function [10]. Note that the density of the Laplace distribution $\text{Lap}(b)$ is given by $f(x) = \frac{1}{2b} \cdot \exp(\frac{-|x|}{b})$, which means that decreasing ϵ will lead to larger expected noise magnitudes, which translates to better privacy.

3 MODEL

We study the problem of designing a set of contracts for *buyers* requesting access to a database managed by a *seller*. We assume that the seller has already acquired data from subjects and compensated them using a one-time monetary payment or a free service (like a phone app). Throughout, we use he/his to refer to buyers and she/her to refer to the seller.

Queries: There are multiple (and finite) types of statistical queries that can be made from the database, denoted by the set $\mathcal{Q}$. The seller offers bundles consisting of a subset of these query types for purchase, with the restriction that any buyer can choose at most one bundle. A bundle is identified by the set $\{Q_1, \ldots, Q_k \mid Q_i \in \mathcal{Q}\}$. The seller designs these bundles based on historical or external information about the types of different buyers, so that every buyers' requirement is met by one of the bundles. Further, for any bundle, the seller limits the number of queries of each type Q_i in the bundle to one (i.e., each bundle is a subset of distinct query types). This follows recommended practices in differential privacy, since allowing multiple queries inevitably degrade privacy guarantees (see also Section 7). We also posit that the seller verifies the identity of buyers, in order to keep track of the buyer's query purchases, and to investigate a privacy attack if it occurs. Further, we posit that the seller, via her service agreement, restricts buyers from faking identifies by imposing substantial post-hoc fines.

Contracts: For each bundle $\{Q_1, \ldots, Q_k\}$, the set of possible contracts are determined by the parameters (p, ϵ, s), with $p \in \mathbb{R}_{\geq 0}$ denoting the price to be paid by the buyer. The privacy levels are assumed to be bounded and normalized such that $\epsilon \in [0, 1]$, with ϵ specifying the bound $\epsilon \geq \epsilon_1 + \ldots + \epsilon_k$, where ϵ_i is used to determine the (Laplace) noise added to the answer of the query of type Q_i; the buyer is free to request any $\epsilon_1, \ldots, \epsilon_k$ within the ϵ bound, with higher ϵ corresponding to less noisy responses. Lastly, s denotes the post-hoc fine to be paid if the buyer is found misusing the query answer.

Buyers: We assume that buyers belong to one of two possible classes: *honest* or *adversarial.*

Honest buyers: Honest buyers do not misuse query answers, and hence generate revenue for the operator when purchasing contracts. Each honest buyer for a given bundle has a type $i \in \Theta := \{1, \ldots, n\}$, determining his benefit from the database. In particular, an honest buyer of type i purchasing contract (p, ϵ, s) derives a *benefit* $b_i(\epsilon) : [0, 1] \longrightarrow \mathbb{R}_{\geq 0}$ from accessing the system. This function includes direct gain from the data, as well as the cost of hedging against the risk of potential direct attack on the buyer. We impose natural conditions on the benefit functions (as is standard for demand functions) $b_i(\cdot)$: that the overall benefit increases with larger ϵ (monotone non-decreasing) and satisfies diminishing returns (concavity), with $b_i(0) = 0$. Most large organizations estimate demand functions and types of buyers from past buyers' activity, and insurance premiums are known; hence, we assume these functions are known. Further, $b_i(\epsilon) \leq b_{i+1}(\epsilon), \forall \epsilon, \forall i$; that is, higher types derive further benefit from the same noise level, e.g., due to their expertise or the relevance of the data to their tasks.

An honest buyer also has a γ probability of suffering an attack himself and causing inadvertent misuse of the query answer, which results in an expected γs loss for him as per the contract terms. Thus, an honest buyer's overall expected utility in its interaction with the seller is given by $u_i(p, \epsilon, s) = b_i(\epsilon) - p - \gamma s$.

Adversarial buyers: An adversarial buyer seeks to access the database with the goal of misusing the information gained. Formally, an adversarial buyer purchasing a bundle through a contract (p, ϵ, s) derives a benefit $C(\epsilon) : [0, 1] \longrightarrow \mathbb{R}_{\geq 0}$ from an attack on the system, with overall adversary utility given by $u_A(p, \epsilon, s) = C(\epsilon) - p - s$. This attack results in a cost $C(\epsilon)$ for the seller. Further, we assume $C(\cdot)$ is monotone increasing and convex, with $C(0) = 0$; intuitively, higher ϵ (lower noise) lead to costlier attacks for the seller, with the severity increasing as the noise decreases. Such convexity has also been noted in literature, e.g., a recent work [16] proposes the cost for seller to be proportional to $\exp(\epsilon) - 1$. Figure 1 shows an example of C and b_i.

We assume that a privacy attack is ultimately discovered, and the seller can track the buyer responsible for the attack. The seller may have to compensate data subjects after a privacy attack (due to lawsuits), which can be partially recovered from the post-hoc fine for data misuse. Note that we have assumed that the adversary cannot cause privacy loss beyond the given ϵ of the bundle by combining the outputs of multiple queries of the same type, as the seller restricts the number of queries per type to one. Further, large post-hoc fines for faking identities prevent the rational adversary from faking identities and attempting to purchase two or more bundles. However, the post-hoc fine for data misuse cannot be set too large as this fine affects the honest buyers, and hence the seller's revenue, due to potential attacks on honest buyers. Therefore, our goal is to study the optimal choice of fines for data misuse so as to deter adversarial buyers while maintaining the demand from honest buyers.

3.1 Seller's revenue optimization problem

We now analyze the seller's contract design problem, with one contract (p, ϵ, s) for each offered bundle. As all rational buyers choose only one bundle due to the marketplace design, these contracts are independent. Therefore, for the rest of this paper, we restrict attention to a given bundle.

Let ρ denote the fraction of adversarial buyers, which is estimated by the seller (conservatively, the seller can estimate ρ to be at most a maximum value).[2] For the honest buyers, let $\{q_i\}_{i\in\Theta}$ denote the fraction of the honest buyers of type i. These fractions can be estimated from historical data. The seller aims to maximize her revenue. Nevertheless, she can not observe individual buyers' types when selling a contract. Consequently, she has to design contracts while balancing two goals: deriving the maximum possible profit from honest types, while limiting the adversarial type's cost to the system.

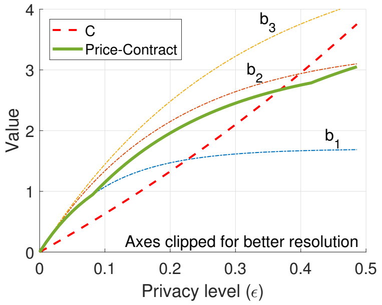

Fig. 1. This figure shows three benefit functions $b_i(\epsilon) = i(1 - \exp(-\frac{10}{i}\epsilon))$ for $i = 1, 2, 3$, an adversary cost function $C(\epsilon) = 6(\exp(\epsilon) - 1)$, and the non-adversarial price-contract curve (defined in Section 4), as a function of the privacy level ϵ.

In classic contract theory, following the *revelation principle* [20, Proposition 14.C.2], it is known that it is enough to offer at most $n+1$ contracts when the number of buyer types is $n+1$.[3] Each agent then selects his intended contract if it satisfies the agent's *individual rationality* (IR) and *incentive compatibility* (IC) constraints. The IR constraint requires that the agent attains higher utility from purchasing the contract compared to opting out. The IC constraint imposes the condition that an agent of type i prefers his intended contract over that of any type $j \neq i$.

Formally, for our contract design problem, consider the $n+1$ user types consisting of the n honest buyers and the adversarial type. Let the contract of type $i \in \Theta$ be (p_i, ϵ_i, s_i) and that for the adversary be (p_A, ϵ_A, s_A). Assume, wlog, that the utility of opting out of purchasing contracts is zero. Then, the IR constraint of an honest buyer of type i, denoted by IR_i, is given by $u_i(p_i, \epsilon_i, s_i) \geq 0$. Similarly, IR_A is $u_A(p_A, \epsilon_A, s_A) \geq 0$. Type i's IC constraints are given by $u_i(p_i, \epsilon_i, s_i) \geq u_i(p_j, \epsilon_j, s_j), \forall j \neq i$ where the j^{th} constraint is denoted by $IC_{i,j}$ and $u_i(p_i, \epsilon_i, s_i) \geq u_i(p_A, \epsilon_A, s_A)$ which is denoted as $IC_{i,A}$. Similarly, the $IC_{A,i}$ constraints can be defined for the adversary.

The seller obtains an expected $p_i + \gamma s_i$ revenue from honest buyers and a revenue of $K(p_A + s_A - C(\epsilon_A))$ from the adversary for some $K > 0$. For ease of presentation, we will fix $K = 1$ but all our results hold for any $K > 0$. The seller's goal is to maximize her overall expected revenue

$$R\big((p_i, \epsilon_i, s_i)_{i\in\Theta}, p_A, \epsilon_A, s_A\big) = (1 - \rho)\Big(\sum_{i=1}^{n} q_i(p_i + \gamma s_i)\Big) + \rho(p_A + s_A - C(\epsilon_A)).$$

However, the seller only has steady (deterministic) revenue over time from p_i; γs_i provides randomly varying revenue over time; in the short term, the realized revenue from fines could be zero as the probability γ is small. Thus, we impose the practical constraint that $p_i \geq (1-\phi)(p_i + \gamma s_i)$, which says that a large fraction $1 - \phi$ of revenue arrive steadily over time. We name this the *steady*

[2] In particular, we postulate that past data misuse and privacy breaches will become known to the seller, by, for instance, directly through inspecting logs of data queries or indirectly through discovery of data misuse by the data subject.
[3] Depending on the type distribution, it may be optimal to offer the same contract to adjacent types (*pooling* contracts).

revenue SR_i constraint. Therefore, the seller's contract design problem can be formally stated as:

$$\begin{aligned} \max_{(p_i,\epsilon_i,s_i)_{i\in\Theta},p_A,\epsilon_A,s_A} \quad & R\big((p_i,\epsilon_i,s_i)_{i\in\Theta},p_A,\epsilon_A,s_A\big) \\ \text{subject to} \quad & IR_i, SR_i\ \forall i \ \text{ and } \ IC_{i,j}\ \forall i,j \text{ and} \\ & IR_A \ \text{ and } IC_{i,A}, IC_{A,i}\ \forall i \text{ and} \\ & p_i,\epsilon_i,s_i \geq 0\ \forall i \text{ and } p_A,\epsilon_A,s_A \geq 0 \end{aligned}$$

3.2 No need for an adversary-specific contract

The contract design problem above includes a contract (p_A, ϵ_A, s_A) for the adversary. While the formulation is mathematically sound and consistent with the revelation principle, this seems an odd design choice as the adversary reveals his type just by choosing this contract. We show that, as intuitively expected, it is in fact not required for the seller to design an adversary-specific contract. That is, despite the fact that the seller faces $n+1$ types, it is optimal to offer at most n contracts.

Lemma 3.1. *The seller should offer at most n contracts/bundles. In particular, it is never optimal to offer an adversary-specific contract/bundle.*

Proof. We show this by contradiction. Assume the seller treats the adversarial buyer as the $(n+1)$-th type, and offers a contract (p_A, ϵ_A, s_A) satisfying all (honest and adversarial) buyers' IR and IC constraints. By IR_A, this contract satisfies $C(\epsilon_A) - p_A - s_A \geq 0$; that is, it will impose a loss $p_A + s_A - C(l\epsilon_A) \leq 0$ on the seller's revenue. Further, by the $IC_{A,i}$ constraints, $C(\epsilon_A) - p_A - s_A \geq C(\epsilon_i) - p_i - s_i$; that is, had the adversary purchased any of the legitimate buyers' contracts, he would have imposed a smaller cost on the seller's revenue. As the seller is a profit-maximizer, we conclude that such contract (p_A, ϵ_A, s_A) should not be part of an optimal collection of contracts. □

Given the above lemma, the contract design problem in the adversarial setting is to design contracts $(p_i, \epsilon_i, s_i)_{i\in\Theta}$ in order to maximize the revenue of the operator:

$$(1-\rho)\Big(\sum_{i=1}^{n} q_i(p_i + \gamma s_i)\Big) + \rho(p_Z + s_Z - C(\epsilon_Z)),$$

where $Z \in \{0, 1, \ldots, n\}$ is the contract chosen by the adversary, subject to IR and IC constraints for all honest buyers in choosing their contract i and the adversary in choosing Z. For the special case of the adversary not choosing any contract, we designate $Z = 0$ with $p_0 = s_0 = \epsilon_0 = 0$. Observe that Z is a variable, and thus, the revenue maximizing problem is a bi-level optimization problem. However, following the standard technique of introducing an additional variable to formulate a zero-sum problem as a linear program, we formulate the revenue maximization problem in the adversarial setting using variable r_A as follows:

$$\begin{aligned} \max_{(p_i,\epsilon_i,s_i)_{i\in\Theta},r_A} \quad & (1-\rho)\big(\textstyle\sum_{i=1}^{n} q_i(p_i+\gamma s_i)\big) + \rho(-r_A) \\ \text{subject to} \quad & IR_i, SR_i\ \forall i \ \text{ and } \ IC_{i,j}\ \forall i,j \text{ and} \\ & r_A \geq C(\epsilon_i) - p_i - s_i\ \forall i \ \text{ and} \\ & p_i,\epsilon_i,s_i \geq 0\ \forall i \ \text{ and } \ r_A \geq 0 \end{aligned}$$

For our described marketplace, one can further consider the corresponding non-adversarial setting, in which the seller solves the contract design problem in the absence of any adversarial considerations. This non-adversarial contract design problem is given by:

$$\begin{aligned} \max_{(p_i,\epsilon_i,s_i)_{i\in\Theta}} \quad & \textstyle\sum_{i=1}^{n} q_i(p_i+\gamma s_i) \\ \text{subject to} \quad & IR_i, SR_i\ \forall i,\ IC_{i,j}\ \forall i,j, \text{ and}, p_i,\epsilon_i,s_i \geq 0\ \forall i \end{aligned}$$

We next study these two contract design problems to characterize the effects of the presence of adversarial types on the optimal contracts' properties and the seller's revenue.

4 ANALYSIS OF ADVERSARIAL CONTRACTING

In classic contract theory, when solving for the optimal contracts, the functions b_i are often assumed to satisfy a condition known as the *single crossing property (SCP)*, which in turn implies the strict increasing differences (ID) property. Throughout our analysis, we will only require the (weaker) condition of (non-strict) ID property on the benefit functions b_i, as defined below:

Definition 4.1 (Increasing Differences). The functions b_i satisfy the (strict) increasing differences property if for any $\epsilon' > \epsilon$, $b_i(\epsilon') - b_i(\epsilon)$ is (strictly) increasing in the type i.

The above condition is a natural assumption on demand functions, and has been used extensively in the contract theory literature starting from the seminal work by [21]. The b_i functions shown in Figure 1 satisfy ID. This condition also allows for significant simplification of the classical contract theory optimization problem. Our first, somewhat surprising result is that, even in the adversarial contract regime with post-hoc fines, the contracts will satisfy a set of constraints akin to those of non-adversarial settings.

Theorem 4.2. *Assuming that the functions b_i satisfy ID, the optimal contracts (in the presence of adversarial types) $(p_1^*, \epsilon_1^*, s_1^*), \ldots, (p_n^*, \epsilon_n^*, s_n^*)$ satisfy the following:*

(1) Monotonicity: $\epsilon_{i+1}^* \geq \epsilon_i^*, \forall i$.
(2) Constraint set reduction: IR_i for $i > 1$ and $IC_{i,j}$ for $j \neq i-1$ are redundant at the optimal contracts.
(3) IR_1 is tight: as a result, $p_1^* + \gamma s_1^* = b_1(\epsilon_1^*)$.
(4) $IC_{i+1,i}$ is tight for all i: as a result for $i > 1$,

$$p_i^* + \gamma s_i^* = b_i(\epsilon_i^*) - \sum_{j=1}^{i-1} \left(b_{j+1}(\epsilon_j^*) - b_j(\epsilon_j^*)\right) .$$

Proof Sketch. As the full proof is rather long we provide a summary here. We first establish the monotonicity of noise levels at the optimal contracts using the (non-strict) ID condition of the benefit functions. Next, we show how to considerably refine the constraint set (point 2) and derive the price-benefit relations (points 3-4). These arguments are based on contradiction: had any of these constraints not been redundant/tight, the operator would have had room to improve her profit by modifying the contracts without violating the remaining IR and IC constraints of honest buyers. For the contradiction argument to carry through, we show that under appropriate modifications, the effect of changes in the adversarial types' behavior on the revenue is non-decreasing. □

Non-adversarial case: We note that for the non-adversarial case, the same results of the above theorem holds; this follows from prior work in contract theory [21] (using a straightforward mapping that we present in the appendix). Formally:

Proposition 4.3. *Assuming that the functions b_i satisfy ID, the optimal contracts in the non-adversarial setting have $s_i^* = 0$ and satisfy all conditions of Theorem 4.2 (with $s_i^* = 0$).*

In particular, the relation between prices, fines, and benefit functions (points 3-4), provides an easy visual representation of the contracts as shown in Figure 1 for the non-adversarial setting (that is, with $s_i = 0$). We call this curve the *price-contract* curve $\mathcal{P}(\epsilon)$, which is a curve plotting the contract prices p (on the y-axis) relative to the privacy levels ϵ (on the x-axis). Specifically, the curve connects the non-adversarial contract points (ϵ_i^*, p_i^*) for all i. From Proposition 4.3, we know that the optimal prices in the non-adversarial setting are such that $p_i^* - p_{i-1}^* = b_i(\epsilon_i^*) - b_i(\epsilon_{i-1}^*)$; thus, the segment of the curve $\mathcal{P}$ that is between ϵ_{i-1}^* and ϵ_i^* is parallel to $b_i(\cdot)$. Therefore, $\mathcal{P}$ is continuous and piece-wise concave. We use $\mathcal{P}$ later in setting up our approximation approach.

Comparison with classic (non-adversarial) contract theory results: Our analysis thus far shows that the design of the optimal contracts in the adversarial and non-adversarial settings carry several similarities. First, it is easy to check that both the adversarial and non-adversarial optimizations are *non-convex* problems which, despite the difference in the number of buyer types, aim to find n optimal contracts. Theorem 4.2's characterization further shows that we can use simplifications similar to those of the non-adversarial contract setting by removing several of the constraints (points 1-2). The result also shows that the optimization problem for computing optimal contracts in the presence of adversaries has only additional adversarial constraints and the same price-benefit relations (points 3-4) as that without adversaries. In particular, the information rent conditions for honest buyers (i.e., the choice of prices that make the honest buyers reveal their types through their contract choice) is the same in both the mixed population and the classic (non-adversarial) setting. Despite these similarities, the presence of adversaries changes the seller's objective function, leading to a different set of contracts than the non-adversarial setting. Proposition 4.3 further implies that the variables s_i can be dropped in the optimization problem for the non-adversarial case, yet these variable remain a key design choice in the adversarial setting.

4.1 Price of Adversary

In order to quantify the effects of the adversary's presence on the seller's revenue, we introduce the following notion:

Definition 4.4 (Price of Adversary). Let R^* and R_A^* denote the seller's maximum revenue in non-adversarial and adversarial settings, respectively. Then, the price of adversary (*PoAdv*) is:

$$PoAdv = (1-\rho)\frac{R^*}{R_A^*}$$

First, note that $PoAdv \geq 1$. This is because the objective of the adversarial optimization problem (i.e. the seller's revenue R_A) is given by $(1-\rho)(\sum_{i=1}^n q_i(p_i + \gamma s_i)) + \rho(-r_A)$, where $r_A \geq 0$ by the problem's constraint set, and the expression in the parenthesis is the revenue attained from honest buyers; therefore, $R_A^* \leq (1-\rho)R^*$. In addition, note that $(1-\rho)$ is included as a normalizing factor in the above definition. This is needed as the measure of honest buyers in the adversarial and non-adversarial cases is different; specifically, it is $(1-\rho)$ and 1, respectively. With the inclusion of this normalizing factor, the smallest value of *PoAdv* will be 1, which is attained when the adversary does not choose any contract, so that $R_A^* = (1-\rho)R^*$. Lastly, if all buyers are adversarial ($\rho = 1$), the revenue attainable by the seller is $R_A^* = 0$, making the *PoAdv* undefined. For this special case, similar to other cases with $R_A^* = 0$, we define $PoAdv = \infty$.

We next analyze the *PoAdv* attainable in the presence of adversaries. Our first finding is that *PoAdv* is unbounded in the worst case (proof is by construction and is presented in the appendix).

Lemma 4.5. *PoAdv is unbounded in the worst case.*

5 APPROXIMATION ALGORITHM

In this section, we present an approach that solves for the adversarial contracting problem approximately, given a solution for the non-adversarial case. We do so since solving the non-adversarial scenario is simpler: by Proposition 4.3, the non-adversarial case has both fewer variables ($s_i = 0, \forall i$) and fewer constraints (no adversary contract choice constraint). Our proposed algorithm also reveals a subtle relation between the adversarial and non-adversarial settings.

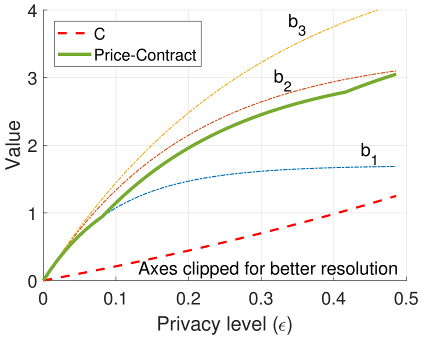

Fig. 2. An example of a *low* C function, reflecting a relatively weak adversary who does not benefit from purchasing any of the offered contracts. Functions $b_i(\cdot), C(\cdot)$ are the same as Fig. 1.

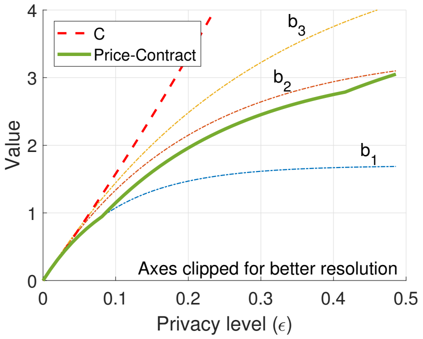

Fig. 3. An example of a *high* C function, reflecting a relatively strong adversary who can benefit from the purchase of any of the offered contracts. Functions $b_i(\cdot), C(\cdot)$ are the same as Fig. 1.

Since by Lemma 4.5 we know that *PoAdv* is unbounded in the worst case, we limit our analysis to a large class of adversary's benefit functions $C(\cdot)$ which imposes mild and natural restriction on these functions. We call these the *well-behaved* C's, and define them as follows. Recall that $\mathcal{P}$ denotes the non-adversarial price-contract curve.

- (Low C) C intersects $\mathcal{P}$ once at the origin and then lies below $\mathcal{P}$ for $\epsilon > 0$; see Figure 2.
- (High C) C intersects $\mathcal{P}$ once at the origin and then lies above $\mathcal{P}$ for $\epsilon > 0$; see Figure 3.
- (Intermediate C) C intersects $\mathcal{P}$ multiple times; see Figure 1. Let $\epsilon_M \in (0, 1)$ be the access level at the last intersection point. We denote $\Delta := \max_{\epsilon < \epsilon_M}\{C(\epsilon) - \mathcal{P}(\epsilon)\}$.

Recall that, by definition, $\mathcal{P}$ connects the optimal non-adversarial contract points (ϵ_i^*, p_i^*). The placement of $C(\cdot)$ relative to these points determines which contracts, if any, yield a positive utility to, and hence would be purchased by, a given adversarial buyer. The above classes therefore comprise several types of adversaries. Low Cs represent weak adversaries who do not find it individually rational to purchase any of the contracts offered to the honest buyers, including the one offered to the lowest type $i = 1$ of honest buyers; this class includes functions $C(\epsilon) \leq b_1(\epsilon)$. In face of such adversaries, no modification to the non-adversarial contracts is needed, hence the term *weak* adversary. High Cs on the other hand represent powerful adversaries, who can afford to impose a high cost on the revenue if the contracts from the non-adversarial setting are offered by the seller, as they would benefit from purchasing any of the contracts, including the one with highest ϵ; this class includes functions $C(\epsilon) \geq b_n(\epsilon)$ as a subset. Finally, intermediate Cs represent adversaries who can purchase (some of) the contracts offered through *non-adversarial* contract design. Within this class, Δ is an upper bound on the adversary's payoff from purchasing contracts with $\epsilon_i^* < \epsilon_M$. As C lies above $\mathcal{P}$ after ϵ_M, we have $C(\epsilon_i^*) \geq p_i^*$ for all $\epsilon_i^* \geq \epsilon_M$, which means that the adversary can afford all contracts with $\epsilon_i^* \geq \epsilon_M$. Figure 1 illustrates an intermediate C. Next, we present our approximation technique. We start with a definition.

Definition 5.1. We call the non-adversarial contract $(p, l, 0)$ a *δ-slack λ-priced* contract, $(\delta, \lambda \geq 0)$, if there exists $s \geq 0$ such that the contract $(p - \gamma s, \epsilon, s)$ satisfies:

- $C(\epsilon) - p - s \leq \delta$, i.e., adversary's gain is bounded by δ.
- $p - \gamma s \geq \lambda > 0$, i.e., the contract's price is at least λ.
- $p - \gamma s \geq (1 - \phi)p$, SR constraint is satisfied

Constructively, s whenever it exists, should be chosen to have the least possible value.

ALGORITHM 1: Approx. Algorithm

```
   Input: Non-adv. contracts $(p_1^*, \epsilon_1^*, 0), \dots, (p_n^*, \epsilon_n^*, 0)$
   Output: An array of contracts or solve adv. case
1  $contracts \leftarrow (p_1^*, \epsilon_1^*, 0), \dots, (p_n^*, \epsilon_n^*, 0)$
2  switch $C$ do
3      case High $C$ do
4          $M = \{i \mid (p_i^*, \epsilon_i^*, 0)$ is 0-slack $\lambda$-priced for some $\lambda \geq 0\}$
5          if $M$ is empty then
6              return solve adv. case
7          $j = \mathrm{argmax}_{k \in M}\, p_k^*$
8          $s_j^* \leftarrow s$ that makes $(p_j^*, \epsilon_j^*, 0)$ 0-slack $\lambda$-priced
9          for $i \leftarrow 1$ to $n$ do
10             $contracts(i) = (p_j^* - \gamma s_j^*, \epsilon_j^*, s_j^*)$
11         return contracts
12     case Low $C$ do
13         return contracts
14     case Intermediate $C$ do
15         return $InterCApp((p_1^*, \epsilon_1^*, 0), \dots, (p_n^*, \epsilon_n^*, 0))$
16 return solve adv. case
```

Using the above definition, our approximation technique is tailored towards the three categories of functions C as shown in Algorithm 1. This algorithm takes the set of non-adversarial contracts as input, and either successfully returns a new set of contracts by modifying this input, or prescribes solving the adversarial contract design problem from scratch. For the High C case, the algorithm finds 0-slack contracts with a positive price (line 4, 0-slack ensures the adversary will not choose the new contract). If one is found, the contract generating the highest revenue among such contracts is offered to all users (line 10). For Low C, the adversary does not choose any contract, hence it is optimal to retain the non-adversarial contracts as is (line 13). For Intermediate C, the function *InterCApp* presented in Algorithm 2 is invoked (line 15).

In Algorithm 2, first a set of Δ-slack p_K^*-priced contracts is found among contracts above and including that of type K (line 2). The best contract with index safe(i) among these is found for each user $i > K$ (line 4). New contracts $(p_{\mathrm{safe}(i)}^* - \gamma s_{\mathrm{safe}(i)}, l_{\mathrm{safe}(i)}^*, s_{\mathrm{safe}(i)})$ are constructed for types $i > K$ (line 6), and all the non-adversarial contracts for types K and below are retained as is (line 8). The revenue from honest buyers for the new contracts is found on line 9, and for the non-adversarial contracts on line 10. β is the utility for the adversarial type in choosing the best new contract (line 11) and α is the same adversary utility in choosing from the non-adversarial contract set (line 12). Line 13-15 compares the revenue in the adversarial setting from the non-adversarial contracts and the new contract set, and returns the contract set that leads to better revenue for the seller.

We next prove that the contracts output by Algorithm 1 are valid. First, we present a lemma on the ordering of honest buyers' preferences over the contracts, which will later be used for the validity proof.

Lemma 5.2. *Given optimal non-adversarial contracts $(p_1^*, \epsilon_1^*, 0), \dots, (p_n^*, \epsilon_n^*, 0)$, a type i user with $i > j$ prefers contract $(p_j^*, \epsilon_j^*, 0)$ over $(p_k^*, \epsilon_k^*, 0)$ for $j > k$.*

The validity of the Algorithm 1's output is as follows:

ALGORITHM 2: *InterCApp*

```
Input: Non-adv. contracts $(p_1^*, \epsilon_1^*, 0), \dots, (p_n^*, \epsilon_n^*, 0)$
Output: An array of contracts
1  $K \longleftarrow$ highest $i$ such that $\epsilon_i^* \le \epsilon_M$                ▷ $\epsilon_M$ as defined in Intermediate $C$
2  $E_{\ge K} = \{k \mid k \ge K$ and $(p_k^*, \epsilon_k^*, 0)$ is $\Delta$-slack $p_K^*$-priced$\}$    ▷ $E_{\ge K}$ not empty as $K \in E_{\ge K}$
3  for $i \longleftarrow K+1$ to $n$ do
4      safe$(i) \longleftarrow \mathrm{argmax}_{k \in E_{\ge K}}\{b_i(\epsilon_k^*) - p_k^*\}$
5      $s^*_{\mathrm{safe}(i)} \longleftarrow s$ that makes $(p^*_{\mathrm{safe}(i)}, \epsilon^*_{\mathrm{safe}(i)}, 0)$ $\Delta$-slack $p_K^*$-priced
6      $contracts(i) = (p^*_{\mathrm{safe}(i)} - \gamma s^*_{\mathrm{safe}(i)}, \epsilon^*_{\mathrm{safe}(i)}, s^*_{\mathrm{safe}(i)})$
7  for $i \longleftarrow 1$ to $K$ do
8      $contracts(i) = (p_i^*, \epsilon_i^*, 0)$
9  $\widehat{R}_K^* = \sum_{i=1}^{K} q_i p_i^* + \sum_{i=K+1}^{n} q_i p^*_{\mathrm{safe}(i)}$
10 $R^* = \sum_{i=1}^{n} q_i p_i^*$
11 $\beta = \max\left(\max_{i \le K}\{C(\epsilon_i^*) - p_i^*\}, \max_{i > K}\{C(\epsilon^*_{\mathrm{safe}(i)}) - p^*_{\mathrm{safe}(i)} - s^*_{\mathrm{safe}(i)}\}\right)$
12 $\alpha = \max_i\{C(\epsilon_i^*) - p_i^*\}$
13 if $(1-\rho)R^* - \rho\alpha > (1-\rho)\widehat{R}_K^* - \rho\beta$ then
14     return $(p_1^*, \epsilon_1^*, 0), \dots, (p_n^*, \epsilon_n^*, 0)$
15 return contracts
```

Lemma 5.3. *For Low or Intermediate Cs, Algorithm 1's output contracts satisfy the IR and IC conditions for all honest buyers. If Algorithm 1 outputs a set of contracts for a High C adversary, then at least one honest buyer buys the contract.*

Proof. For High C, there is one contract offered to all buyers, so the IC constraints are trivially satisfied. Also, for user j the contract offered satisfies IR, since from optimality of the non-adversarial contracts we get $b_j(\epsilon_j^*) - p_j^* \ge 0$. For Low C, the set of non-adversarial contracts are returned by the algorithm, and so the proof is immediate from optimality of the non-adversarial contracts.

For Intermediate C, if the set of non-adversarial contracts are returned by Algorithm 2, then the claim again holds trivially. Otherwise, assume new contracts are returned. Observe that the contract $(p_K^*, \epsilon_K^*, 0)$ is Δ-slack p_K^*-priced (follows from Def. 5.1 and def. of K, Δ). Thus, $E_{\ge K}$ is not empty as $K \in E_{\ge K}$. Also, note that for users $i > K$, the offered modified contracts (line 6) have the effective price $p^*_{\mathrm{safe}(i)} - \gamma s^*_{\mathrm{safe}(i)} + \gamma s^*_{\mathrm{safe}(i)} = p^*_{\mathrm{safe}(i)}$, same as the non-adversarial contract.

We first start by analyzing users $i > K$. Fix i to be any index $> K$. All users $j > K$ (including i) are offered modified contracts (line 6) from among those indexed by $E_{\ge K}$ (loop on line 3). By definition of **safe**(i), $b_i(\epsilon^*_{\mathrm{safe}(i)}) - p^*_{\mathrm{safe}(i)} \ge b_i(\epsilon_k^*) - p_k^*$ for all $k \in E_{\ge K}$. Thus, i prefers his contract over any other offered to any $j > K$. Next, by definition of safe(i), $b_i(\epsilon^*_{\mathrm{safe}(i)}) - p^*_{\mathrm{safe}(i)} \ge b_i(\epsilon_K^*) - p_K^*$, and then by Lemma 5.2 and our case of $i > K$, $b_i(\epsilon_K^*) - p_K^* \ge b_i(\epsilon_j^*) - p_j^*$ for all $j \le K$. Thus, $b_i(\epsilon^*_{\mathrm{safe}(i)}) - p^*_{\mathrm{safe}(i)} \ge b_i(\epsilon_j^*) - p_j^*$ for all $j \le K$ and hence all the IC constraints for i are satisfied. For IR, first by the ID property we have $b_i(\epsilon_K^*) \ge b_K(\epsilon_K^*)$, hence $b_i(\epsilon_K^*) - p_K^* \ge b_K(\epsilon_K^*) - p_K^* \ge 0$, where the ≥ 0 is due to optimality of the non-adversarial contracts. Finally, we just proved that $b_i(\epsilon^*_{\mathrm{safe}(i)}) - p^*_{\mathrm{safe}(i)} \ge b_i(\epsilon_K^*) - p_K^*$, thus, $b_i(\epsilon^*_{\mathrm{safe}(i)}) - p^*_{\mathrm{safe}(i)} \ge 0$.

Next, the users $i \le K$ are offered the non-adversarial contracts, thus, $b_i(\epsilon_i^*) - p_i^* \ge b_i(\epsilon_j^*) - p_j^*$ for all $j \ne i$. Since the modified contracts (line 6) still have an effective price same as the non-adversarial contract, any user $i \le K$ still prefers his contract to the modified ones. The IR constraint is also satisfied as the non-adversarial contracts were optimal. □

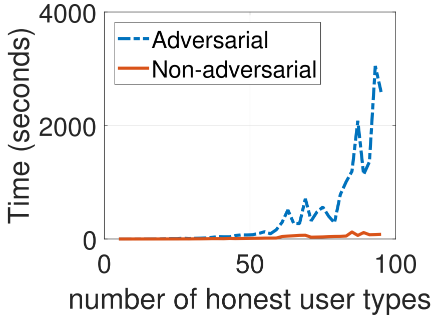

Fig. 4. Runtime comparison

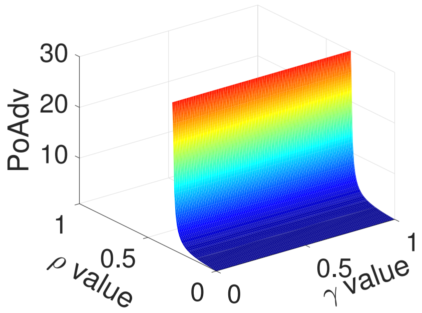

Fig. 5. *PoAdv* for non-adversarial contracts

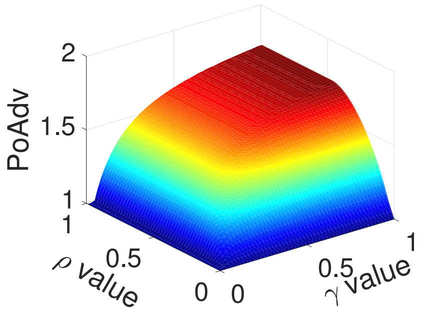

Fig. 6. *PoAdv* for optimal adversarial contracts

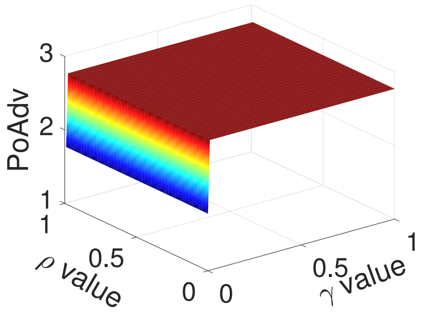

Fig. 7. *PoAdv* for approx adversarial contracts

Next, the following result establishes the quality of the contracts returned by Algorithm 1 by bounding the *PoAdv*. Recall that we have already shown in Lemma 4.5 that *PoAdv* is unbounded in the worst case.

Theorem 5.4. *Let the optimal non-adversarial contract* $(p_1^*, \epsilon_i^*), \dots, (p_n^*, \epsilon_n^*)$ *revenue be* R^*. *For the class of well-behaved* C*'s, we have,*

- *(High* C*) PoAdv is unbounded in general. If Algorithm 1 outputs a contract, then* $PoAdv \leq \frac{R^*}{\lambda \min_i \{q_i\}}$.
- *(Low* C*) Algorithm 1 always outputs the same contracts as the non-adversarial case, and hence* $PoAdv = 1$.
- *(Intermediate* C*) Alg. 1 always outputs contracts. Then,*

$$PoAdv \leq \frac{R^*}{\max\left(\widehat{R}_K^* - \Delta\frac{\rho}{1-\rho}, R^* - \alpha\frac{\rho}{1-\rho}\right)}\,.$$

Proof. For High C, if a contract is offered, the adversary will not choose this contract due to 0-slack, but at least one honest buyer i will choose it. Thus, the revenue is at least $(1-\rho)\lambda \min_i q_i$. For low C, clearly the adversary does not choose any contract, and $R_A^* = (1-\rho)R^*$.

For intermediate C, the revenue from the contracts output by Algorithm 1 is $\widehat{R}_K^*$. The adversary may choose any of the contracts. If the choice is $Z \leq K$ (all of which have fine 0), then by definition of intermediate functions and K, we will have $C(\epsilon_Z^*) - p_Z^* \leq \Delta$ for all such Z. If $Z > K$, then since the offered contracts $> K$ are all Δ-slack, we again have $C(\epsilon_Z^*) - p_Z^* - s_Z^* \leq \Delta$. Thus, $\beta \leq \Delta$, and the revenue is lower bounded by $(1-\rho)\widehat{R}_K^* - \rho\Delta$. Finally, by not changing the original non-adversarial contracts, the operator obtains a revenue $(1-\rho)R^* - \rho\alpha$. Thus, the revenue in the presence of adversaries is bounded by the maximum of either of these two lower bounds. □

The following corollary bounds the approximation performance in terms of approximation ratio.

Corollary 5.5 (Approximation Ratio). *Let B denote the bound on PoAdv on Theorem 5.4 and let U denote the revenue provided by the approximation algorithm. Then, U satisfies $U > (1/B)R_A^*$, that is, U is at least $1/B$ of the optimal adversarial revenue.*

Proof. By definition of *PoAdv* and the bound B in Theorem 5.4, we get $(1-\rho)R^*/U < B$ and also $R_A^* \leq (1-\rho)R^*$. Hence using these two we get $U > (1/B)R_A^*$. □

6 NUMERICAL EXAMPLE

While our theory results provide a broad characterization of the problem for a large space of utility functions, in this section we illustrate specific points related to the problem parameters, with a numerical example. We use $n = 10$ types of honest buyers (except when varying n), with $b_i(\epsilon) = i(1 - \exp(-\frac{10\epsilon}{i}))$, $C(\epsilon) = 6(\exp(\epsilon) - 1)$, and $\phi = 0.95$.

Runtime comparison: Fig. 4 illustrates runtimes for computing the optimal adversarial and optimal non-adversarial contracts. The optimal adversarial contracts take much more time to compute than the non-adversarial contracts and the difference increases exponentially with increase in the size of problem n. This shows why approximation is useful; our approach takes almost the same time as the non-adversarial problem (thus, not shown in Fig. 4), as the approximation steps after solving the non-adversarial problem have (comparatively) negligible runtime.

PoAdv **with non-adversarial contracts**: Fig. 5 shows the *PoAdv* for varying γ and ρ when non-adversarial contracts are offered in an adversarial setting. The *PoAdv* rises sharply with ρ. Intuitively, the non-adversarial contracts suffer great loss if adversarial buyers dominate the market.

PoAdv **with optimal adversarial contracts**: Fig. 6 shows the *PoAdv* for varying γ and ρ when the optimal adversarial contracts are computed exactly. The *PoAdv* rises with both increasing γ and ρ. Intuitively, higher ρ represents adversaries' market domination, and higher γ is weaker honest users (i.e., more attack-prone). Thus, higher values for both of these parameters cause more loss, leading to higher *PoAdv*.

Performance of approximation: Lastly, Fig. 7 shows the *PoAdv* computed using our approximation approach for varying γ and ρ. The C that we chose corresponds to an Intermediate C. The *PoAdv* varies mostly with γ and is almost constant throughout at 2.77, except for very small values of γ when it is 1.43. For small values of γ, the approximation algorithm sends back the original contracts as is (line 14 in Algorithm 2).

7 RELATED WORK AND SUMMARY

Our work is within the emerging literature of data commercialization and its challenges [32]. Both [18] and [32] discuss the profit opportunities from packaging data based on the users' needs and willingness to pay; we formalize these notions through the framework of contract design, with a focus on data privacy preservation.

A number of recent papers have studied the design of optimal pricing mechanisms for data sellers. Specifically, the works of [12, 13, 19] study the problem of pricing personal data, where a data seller designs a pricing mechanism which incentivizes data subjects to reveal their private information. The work of [4] compares the two pricing mechanisms of upfront payments and pay-per-use from the viewpoint of data sellers. The authors of [27] design a pricing scheme for selling data to users with differing willingness to pay. Our approach differs from these works in that we propose a contract-theoretical framework to accommodate heterogeneous honest buyers as well as adversarial types. More specifically, in contrast to existing work, we posit that honest buyers do not attempt to misuse the information gained from the database, hence every sale of data is not a privacy attack. Further, by far the practice in real world is for the data seller to obtain data by compensating people in form of a one-shot monetary payment or free service [28], which is part of our model. This avoids practically unrealizable mechanisms in which data subjects are paid every time their data is sold to a buyer [19]. Recent approaches have also looked at enforcing rules and regulations (such as what we propose) using blockchains [24], including in marketplaces [5]. These approaches complement our economic driven approach by providing a technical rather than legal means of enforcing fines, etc. Other works model problems where the buyer directly buys data from data subjects [11], which is not the problem setting in data marketplaces.

Adam and Worthmann [2] classified privacy-preserving query approaches into query restriction, data perturbation, and output perturbation. Query auditing (a form of query restriction) aims to determine whether, given the query history, a new query will compromise the database privacy; however, this problem is NP-hard [17]. In addition, output perturbation mechanisms (including differential privacy) must limit the number of queries in order to maintain any reasonable privacy guarantee [8]. Our proposed approach, which is a combination of query restriction with output perturbation, restricts the type and number of queries in light of these impossibility results.

Contract-theoretical frameworks have been receiving attention as a method for optimal pricing in other application areas, including the design of demand-response programs [22], energy procurement methods [31], and incentive mechanisms in crowdsourcing markets [15]. In contrast, we consider the optimal pricing problem in the presence of both honest and adversarial buyers.

Another line of work studies the effects of malicious or spiteful agents in game-theoretical settings including auctions such as network inoculation games [25], sealed-bid auctions and colluding bidders [6, 23], and resource allocation games [7]. These works assume that malicious agents aim to minimize the utility of *all other users*, and analyzes their effect on the Nash equilibria. In contrast, we consider the effects of an adversarial user on the principal's revenue. Other work consider privacy concerns in revealing agent's types [26] or complexity of contracts in the hidden action setting [3], which are quite distinct from our focus in this paper.

Summary: We proposed a novel and practical *adversarial contract design* framework in which a data seller designs a collection of contracts to optimize her revenue in the presence of honest and adversarial buyers. We proposed that the seller add noise to data query answers, charge more for lower noise, and thwart rational adversaries by levying fines. We quantified the effect of adversaries by proposing the price of adversary, and characterized the effect of fines on optimal revenue. Finally, we presented a fast approximate technique to compute contracts in the adversarial setting.

Acknowledgement. This work was supported in part by the MURI grant W911NF-18-1-0208 from the US Army Research Office.

REFERENCES

[1] Acxiom. 2018. Acxiom. https://www.acxiom.com/how-we-can-help/data-stewardship/. Accessed: 2018-12-15.

[2] Nabil R Adam and John C Worthmann. 1989. Security-control methods for statistical databases: a comparative study. *ACM Computing Surveys (CSUR)* 21, 4 (1989), 515–556.

[3] Moshe Babaioff and Eyal Winter. 2014. Contract complexity. In *Proceedings of the 15th ACM Conference on Economics and Computation.*

[4] Sridhar Balasubramanian, Shantanu Bhattacharya, and Vish V Krishnan. 2015. Pricing information goods: A strategic analysis of the selling and pay-per-use mechanisms. *Marketing Science* 34, 2 (2015), 218–234.

[5] Prabal Banerjee and Sushmita Ruj. 2018. Blockchain Enabled Data Marketplace - Design and Challenges. *CoRR* abs/1811.11462 (2018). arXiv:1811.11462 http://arxiv.org/abs/1811.11462

[6] Felix Brandt, Tuomas Sandholm, and Yoav Shoham. 2007. Spiteful Bidding in Sealed-Bid Auctions.. In *IJCAI*, Vol. 7. 1207–1214.

[7] Anil Kumar Chorppath and Tansu Alpcan. 2011. Adversarial behavior in network mechanism design. In *Proceedings of the 5th International ICST Conference on Performance Evaluation Methodologies and Tools*. ICST, 506–514.

[8] Irit Dinur and Kobbi Nissim. 2003. Revealing information while preserving privacy. In *Proceedings of the 22nd ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. ACM, 202–210.

[9] Cynthia Dwork. 2008. Differential privacy: A survey of results. In *International Conference on Theory and Applications of Models of Computation*. Springer, 1–19.

[10] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*. Springer, 265–284.

[11] Arpita Ghosh, Katrina Ligett, Aaron Roth, and Grant Schoenebeck. 2014. Buying private data without verification. In *Proceedings of the fifteenth ACM conference on Economics and computation*. ACM, 931–948.

[12] Arpita Ghosh and Aaron Roth. 2011. Selling Privacy at Auction. In *Proceedings of the 12th ACM Conference on Electronic Commerce (EC '11)*. 199–208.

[13] Vasilis Gkatzelis, Christina Aperjis, and Bernardo A Huberman. 2015. Pricing private data. *Electronic Markets* 25, 2 (2015), 109–123.

[14] Kevin Granville. 2018. Facebook and Cambridge Analytica: What You Need to Know as Fallout Widens. https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html.

[15] Chien-Ju Ho, Aleksandrs Slivkins, and Jennifer Wortman Vaughan. 2016. Adaptive contract design for crowdsourcing markets: Bandit algorithms for repeated principal-agent problems. *Journal of Artificial Intelligence Research* 55 (2016), 317–359.

[16] Justin Hsu, Marco Gaboardi, Andreas Haeberlen, Sanjeev Khanna, Arjun Narayan, Benjamin C Pierce, and Aaron Roth. 2014. Differential privacy: An economic method for choosing epsilon. In *Computer Security Foundations Symposium (CSF), 2014 IEEE 27th*. IEEE, 398–410.

[17] Jon Kleinberg, Christos Papadimitriou, and Prabhakar Raghavan. 2003. Auditing boolean attributes. *J. Comput. System Sci.* 66, 1 (2003), 244–253.

[18] Pantelis Koutroumpis and Aija Leiponen. 2013. Understanding the value of (big) data. In *The 2013 IEEE International Conference on Big Data*. IEEE, 38–42.

[19] Chao Li, Daniel Yang Li, Gerome Miklau, and Dan Suciu. 2014. A theory of pricing private data. *ACM Transactions on Database Systems (TODS)* 39, 4 (2014), 34.

[20] Andreu Mas-Colell, Michael Dennis Whinston, Jerry R Green, et al. 1995. *Microeconomic theory*. Vol. 1. Oxford university press New York.

[21] Eric Maskin and John Riley. 1984. Monopoly with incomplete information. *The RAND Journal of Economics* 15, 2 (1984), 171–196.

[22] Reshef Meir, Hongyao Ma, and Valentin Robu. 2017. Contract design for energy demand response. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*. AAAI Press, 1202–1208.

[23] S. Micali and P. Valiant. 2008. *Revenue in Truly Combinatorial Auctions and Adversarial Mechanism Design*. Technical Report MIT-CSAIL-TR-2008-039. MIT.

[24] Carlos Molina-Jimenez, Ioannis Sfyrakis, Ellis Solaiman, Irene Ng, Meng Weng Wong, Alexis Chun, and Jon Crowcroft. 2018. Implementation of Smart Contracts Using Hybrid Architectures with On and Off–Blockchain Components. In *2018 IEEE 8th International Symposium on Cloud and Service Computing (SC2)*. IEEE, 83–90.

[25] Thomas Moscibroda, Stefan Schmid, and Rogert Wattenhofer. 2006. When selfish meets evil: Byzantine players in a virus inoculation game. In *Proceedings of the 25th annual ACM symposium on Principles of distributed computing*. 35–44.

[26] Kobbi Nissim, Claudio Orlandi, and Rann Smorodinsky. 2012. Privacy-aware mechanism design. In *Proceedings of the 13th ACM Conference on Electronic Commerce*. ACM, 774–789.

[27] Dusit Niyato, Mohammad Abu Alsheikh, Ping Wang, Dong In Kim, and Zhu Han. 2016. Market model and optimal pricing scheme of big data and Internet of things (IoT). In *The 2016 IEEE International Conference on Communications*.

[28] Eduardo Porter. 2018. Your Data Is Crucial to a Robotic Age. Shouldn't You Be Paid for It? https://www.nytimes.com/2018/03/06/business/economy/user-data-pay.html. Accessed: 2018-12-15.

[29] salesforce.com, inc. 2018. Salesforce DMP Security, Privacy and Architecture. https://help.salesforce.com/servlet/servlet.FileDownload?file=0150M0000041PNOQA2. Accessed: 2018-12-15.

[30] Sarah Spiekermann, Rainer Böhme, Alessandro Acquisti, and Kai-Lung Hui. 2015. Personal data markets. *Electronic Markets* 25, 2 (2015), 91–93.

[31] Hamidreza Tavafoghi and Demosthenis Teneketzis. 2014. Optimal contract design for energy procurement. In *Communication, Control, and Computing (Allerton), 2014 52nd Annual Allerton Conference on*. IEEE, 62–69.

[32] L. D. W. Thomas and A. Leiponen. 2016. Big data commercialization. *IEEE Engineering Management Review* 44, 2 (Second 2016), 74–90. https://doi.org/10.1109/EMR.2016.2568798

A APPENDIX: OMITTED PROOFS

Proof of Theorem 4.2. As a shorthand, we will write $p'_i = p_i + \gamma s_i$ throughout.

Monotonicity: First, we claim that for every optimal fixed cost contract we must have $\epsilon_i \geq \epsilon_j$ whenever $i > j$. Let $i > j$. The IC constraints include

$$b_i(\epsilon_i) - p'_i \geq b_i(\epsilon_j) - p'_j \text{ and } b_j(\epsilon_j) - p'_j \geq b_j(\epsilon_i) - p'_i$$

Adding these, we get

$$b_i(\epsilon_i) - b_i(\epsilon_j) \geq b_j(\epsilon_i) - b_j(\epsilon_j)$$

There are two cases (1) $b_i(\epsilon_i) - b_i(\epsilon_j) > b_j(\epsilon_i) - b_j(\epsilon_j)$ or (2) $b_i(\epsilon_i) - b_i(\epsilon_j) = b_j(\epsilon_i) - b_j(\epsilon_j)$. For case (1), we can claim that $\epsilon_i \geq \epsilon_j$ using non-strict ID of the benefits functions. The proof is by contradiction. Assume $\epsilon_i < \epsilon_j$; then, by non-strict ID we must have $b_i(\epsilon_i) - b_i(\epsilon_j) \leq b_j(\epsilon_i) - b_j(\epsilon_j)$ which violates case (1). Hence under case (1) $\epsilon_i \geq \epsilon_j$. As the reasoning here is not based on the seller's objective value or the adversarial type's constraints, we do not need to consider adversarial aspects here.

Next, under case (2), let $K = b_i(\epsilon_i) - b_i(\epsilon_j) = b_j(\epsilon_i) - b_j(\epsilon_j)$. First, if K is ≥ 0 then $\epsilon_i \geq \epsilon_j$ when b_i and b_j are both strictly monotone increasing. As the reasoning here is not based on the objective value or the adversarial constraints, we do not need to consider adversarial aspects here. The case when b_i and b_j are both monotone non-decreasing has to be dealt in a special way (see after the $K < 0$ case below).

Thus, the only scenario left to analyze is $K < 0$. Then the two IC inequalities stated at the start can be re-written as $K \geq p'_i - p'_j$ and $-K \geq -(p'_i - p'_j)$ which implies $p'_i - p'_j = K$, or $p'_i < p'_j$. Also, $b_i(\epsilon_i) - p'_i = b_i(\epsilon_j) + K - p'_j - K = b_i(\epsilon_j) - p'_j$, so that contract i and j are both equally and *most* preferred by i (and similarly by j). Then offer another set of contracts in which i is offered (p_j, ϵ_j, s_j), and others are offered their earlier contract. In this new contract, all of the IC constraints are still satisfied as type i preferred (p_j, ϵ_j, s_j) the most and equally preferred the now unavailable (p_i, ϵ_i, s_i). For any other type they prefer their allocation and price to (p_j, ϵ_j, s_j) as was the case for the earlier set of contracts. Also, since $b_i(\epsilon_i) - p'_i = b_i(\epsilon_j) - p'_j$ and earlier contract's IR provided $b_j(\epsilon_i) - p'_i \geq 0$, we have the new contract's IR is also satisfied $b_i(\epsilon_j) - p'_j \geq 0$. The SR_i is also trivially satisfied since SR_j was satisfied. In this new set of contracts, as $p'_j > p'_i$ the revenue from i increases and all other honest users provide same revenue as earlier, thus, the operator's revenue from the honest users strictly increases. Finally, we need to analyze the adversaries incentives in this new collection of contracts. For the adversary, the new set of contracts provides fewer options to choose from; thus, for any choice made by the adversary in the new contract regime, he obtains less or equal utility to that from the original contract set. As the operator's utility is zero-sum with the adversary's utility, the contribution from the adversarial part of the operator's revenue either increases or stays the same in the new set of contracts. Thus, putting these together, we have found a new, feasible set of contracts, that strictly outperforms the original set of contracts, contradicting the optimality of the original set. Hence, we cannot have $K < 0$.

Special case (non-decreasing b_i and b_j): The case when b_i and b_j are both monotone non-decreasing requires to treat the special case of $K = 0$ separately. Thus, reasoning exactly like the $K < 0$ case we get that $p'_i = p'_j$ and $b_i(\epsilon_i) = b_i(\epsilon_j)$ and (p_i, ϵ_i, s_i) and (p_j, ϵ_j, s_j) are both equally preferred by i. Now, if $\epsilon_i \geq \epsilon_j$ we are done, but if not we can offer (p_j, ϵ_j, s_j) to i. Following an argument similar to the case of $K < 0$, the new set of contracts would satisfy all IR, SR, and IC constraints of the honest types. From the seller's viewpoint, the overall revenue from legitimate users remains the same as the original set of contracts. Further, following an argument similar to case $K < 0$, the contribution from the adversarial part of the revenue either increases or stays the same with the new set of contracts. Therefore, for this special case, we can claim that if $\epsilon_i < \epsilon_j$, the set of contracts is revenue equivalent (or even suboptimal to) a collection of contracts with $\epsilon_i = \epsilon_j$. We conclude that at the optimal contract, $\epsilon_i \geq \epsilon_j$ for this case as well.

Constraint-set reduction: Next, we move on to the IC and IR constraints' properties. We start with the IR constraints. Starting from IR_i, we have,

$$b_i(\epsilon_i) - p'_i \geq b_i(\epsilon_{i-1}) - p'_{i-1} \geq b_{i-1}(\epsilon_{i-1}) - p'_{i-1} \ ,$$

where the first line follows from $IC_{i,i-1}$, and the second line by the assumption on ordering of the benefit functions, i.e., $b_i(l) \geq b_j(l), \forall i > j, \forall l$. Thus, if IR_{i-1} is satisfied so is IR_i. Hence, given IR_1 is satisfied, all other IR constraints are redundant. As the reasoning here is not based on objective value or adversary constraints, this assertion holds both with and without adversarial types.

Next, we consider the (IC) constraints. By $IC_{i-1,i-2}$ we have $b_{i-1}(\epsilon_{i-1}) - p'_{i-1} \geq b_{i-1}(\epsilon_{i-2}) - p'_{i-2}$, which can be rearranged as $b_{i-1}(\epsilon_{i-1}) - b_{i-1}(\epsilon_{i-2}) \geq p'_{i-1} - p'_{i-2}$. By non-strict increasing difference and, as shown earlier, the monotonicity of access levels, $\epsilon_{i-1} \geq \epsilon_{i-2}$, we get,

$$b_i(\epsilon_{i-1}) - b_i(\epsilon_{i-2}) \geq b_{i-1}(\epsilon_{i-1}) - b_{i-1}(\epsilon_{i-2}) \geq p'_{i-1} - p'_{i-2}.$$

Thus, $b_i(\epsilon_{i-1}) - p'_{i-1} \geq b_i(\epsilon_{i-2}) - p'_{i-2}$. By $IC_{i,i-1}$, we have $b_i(\epsilon_i) - p'_i \geq b_i(\epsilon_{i-1}) - p'_{1-i}$, and hence we can infer that $b_i(\epsilon_i) - p'_i \geq b_i(\epsilon_{i-2}) - p'_{i-2}$. Thus, given the local downward IC constraints $IC_{i-1,i-2}$ and $IC_{i,i-1}$, the $IC_{i,i-2}$ constraint is redundant; similarly, all $IC_{i,i-k}$ constraints are redundant for $k \geq 2$. Next, for the local upward IC constraints, starting from $IC_{i+1,i+2}$, we have $b_{i+1}(\epsilon_{i+1}) - p'_{i+1} \geq b_{i+1}(l_{i+2}) - p'_{i+2}$, which can be rearranged as $p'_{i+2} - p'_{i+1} \geq b_{i+1}(l_{i+2}) - b_{i+1}(\epsilon_{i+1})$. Again, by non-strict increasing difference and monotonicity $\epsilon_{i+1} \geq l_i$, we'll get $b_i(\epsilon_{i+1}) - p'_{i+1} \geq b_i(l_{i+2}) - p'_{i+2}$. Thus, we conclude that given the local upward IC constraints, all other upward constraints $IC_{i,i+k}$ for $k \geq 2$ are redundant. Hence, only the local $IC_{i,i+1}$ and $IC_{i,i-1}$ constraints are non-redundant. As the reasoning here is not based on objective value or adversary constraints, the arguments remain valid in the presence of adversaries.

Next, we show that the local upward IC constraints $IC_{i,i+1}$ is also redundant. For contradiction, suppose we solve the optimization problem without the $IC_{i,i+1}$ constraint, and get the set of contracts $\{p_j, \epsilon_j, s_j\}$ that maximize the operator's revenue. This solution should strictly violate $IC_{i,i+1}$ (since we are assuming $IC_{i,i+i}$ is not redundant). Therefore, type i will strictly prefer the contract $\{p_{i+1}, \epsilon_{i+1}, s_{i+1}\}$, that is, $b_i(\epsilon_{i+1}) - b_i(\epsilon_i) > p'_{i+1} - p'_i$. We now modify the contracts by increasing $p_j, \forall j \geq i+1$ by a small amount $\epsilon > 0$, i.e., we offer the contract $\{p_{i+1} + \epsilon, \epsilon_{i+1}, s_{i+1}\}$ to type $i+1$, as well as contracts $\{p_j + \epsilon, \epsilon_j, s_j\}$ for all $j > i+1$. We chose ϵ small enough so that $IC_{i,i+1}$ remains strictly violated.

We know from the violation of $IC_{i,i+1}$ that $b_i(\epsilon_{i+1}) - b_i(\epsilon_i) > p'_{i+1} - p'_i$, and also, by non-strict increasing differences, that $b_{i+1}(\epsilon_{i+1}) - b_{i+1}(\epsilon_i) > p'_{i+1} - p'_i$, or rearranging $b_{i+1}(\epsilon_{i+1}) - p'_{i+1} > b_{i+1}(\epsilon_i) - p'_i$; thus, $IC_{i+1,i}$ is satisfied with $\{p_{i+1} + \epsilon, \epsilon_i, s_i\}$. For all other local upward IC constraints of types $i+1$ and higher (i.e, $IC_{i+1,i+2}$, $IC_{i+2,i+1}$, $IC_{i+2,i+3}$ and so on), the prices on both sides of the constraint change by an equal amount in the modified contract set. Therefore, these constraints continue to hold. For all other IC constraints there is no change in variable values and they continue to hold. The IR_1 constraint is also unaffected as the contract does not change for type 1. All SR constraints still hold as only prices increased. For the adversary, the contracts in the new collection are either the same (if he was purchasing one of the unaltered contracts) or become less attractive (if he was purchasing the altered contract). Thus, for any choice made by the adversary in the new contract regime, he obtains either less or the same utility as the original contract set. As the seller's and adversary's utilities are zero-sum, the contribution from the adversarial part of the revenue either increases or stays the same following the change in the contracts. Thus, this new set of contracts provides higher revenue to the operator, contradicting the optimality of the original set of contracts. We conclude that all local upward IC constraints should be redundant.

IR_1 **is redundant:** we prove this by contradiction. Suppose IR_1 is not binding; then, the operator can increase p_1 slightly without violating IR_1 (and trivially not violating SR_1). The only other constraint in which in which p_1 appears is the LHS of the downward IC constraint IC_{21}. An increase in p_1 will lower the LHS, and hence this constraint will not be violated either. Therefore, the operator's portion of the revenue from legitimate users is strictly increasing with this increase in p_1. From the adversary's viewpoint, the new set of contracts (with an increased p_1 in the lowest type's contract) will either stay the same or becomes less attractive. Thus, for any choice made by the adversary in the new contract regime, he obtains less or equal utility to his utility in the original contract set. As the seller's revenue portion from the adversarial type's participation is the negative of the adversary's utility, the contribution from the adversarial part of the revenue will either increase or stay the same given the increase in p_1. Thus, the modification of the price p_1 will lead to a feasible set of contracts that strictly increases the operator's revenue, contradicting the optimality of the original contract set. We thus conclude that IR_1 should be binding in the optimal contract set, so that $p_1^* + \gamma s_1^* = b_1(\epsilon_1^*)$.

($IC_{i,i-1}$) **is binding:** finally, we show that all the $IC_{i,i-1}, \forall i \geq 2$ constraints are binding. For contradiction, suppose $IC_{i,i-1}$ is not binding. Then we can increase p_i by ϵ without violating this constraint. In all remaining local downward IC constraint, p_i only appears on the LHS of $IC_{i+1,i}$; the increase in p_i will therefore not violate this constraint. In addition, IR_1 will not be affected and also SR_i constraint will not be violated as p_i only appears on the LHS of SR_i, and the revenue of the operator from legitimate users will strictly increase following this change. For the adversary, for all contracts in the new set of contract, the contract either stays the same or becomes less attractive for the adversary (due to higher price). Thus, for any choice made by the adversary in the new contract regime, he obtains less or equal utility to that from the original contract set. As the seller receives the negative of the adversary's utility, the contribution from the adversarial part of the objective either increases or stays the same, and hence the overall revenue of the operator increases with the modified contract set. This provides a contradiction to the optimality of the initial contracts. Therefore, the local downward IC constraints should be binding, leading to,

$$p_{i+1}^* + \gamma s_{i+1}^* = b_{i+1}(l_{i+1}^*) - \sum_{j=1}^{i} \left(b_{j+1}(\epsilon_j^*) - b_j(\epsilon_j^*) \right) .$$

□

Proof of Proposition 4.3. First, we will prove that for any optimal solution with non-zero s_i's there is a revenue (objective) equivalent solution with all s_i zero. The transformation is simple: given any solution (p_i, ϵ_i, s_i), the contract $(p_i + \gamma s, \epsilon_i, 0)$ is feasible and revenue optimal. The revenue stays the same, which trivially follows from the objective function. All constraints, except SR, have the term $p_i + \gamma s$, and hence they are satisfied. The SR constraints are trivially satisfied as s_i is zero in the new contract. Next, with contracts for which s_i is 0, the optimization reduces to

$$\begin{array}{ll} \max\limits_{(p_i,\epsilon_i)_{i\in\Theta}} & \sum_{i=1}^{n} q_i p_i \\ \text{subject to} & IR_i \ \forall i \text{ and } IC_{i,j} \ \forall i,j \text{ and } p_i, \epsilon_i \geq 0 \ \forall i \end{array}$$

This is exactly same as the classic contract theory problem, and the conditions of Theorem 4.2 (with $s_i^* = 0$) follow from the seminal work by Maskin and Riley [21] □

Proof of Lemma 4.5. Consider a problem with two types of honest buyers H and L. Let the benefit function be $\log(1 + \epsilon)$ for the lower type L and $2\log(1 + \epsilon)$ for the higher type H. The buyer is type L with probability q and H type with probability $1 - q$. The function C for adversary is $K(\exp(\epsilon) - 1)$, where K will be chosen below. For now, let $K \geq 2 + 2(1 - \gamma)/\gamma$.

For the adversarial revenue maximization case we will show the revenue is 0. Let the contract with the adversary with (p_L, ϵ_L, s_L) and (p_H, ϵ_H, s_H). To show 0 revenue we will show that $\epsilon_H = 0$ (and since $\epsilon_L \leq \epsilon_H$, $\epsilon_L = 0$). We do so by contradiction, whereby assume $\epsilon_H > 0$. First, it directly follows from Theorem 4.2 that $p_L + \gamma s_L = \log(1 + \epsilon_L)$ and $p_H + \gamma s_H \leq 2\log(1 + \epsilon_H)$. Next, as $\epsilon_H \geq \epsilon_L$, we have $p_L + \gamma s_L \leq \log(1 + \epsilon_H)$.

The adversary never rejects the higher contract, since for any $\epsilon_H \in (0, 1]$,

$$\begin{aligned} K(\exp(\epsilon_H) - 1) &\geq (2 + 2(1-\gamma)/\gamma)(\exp(\epsilon_H) - 1) \\ &\geq (2 + 2(1-\gamma)/\gamma)\log(1 + \epsilon_H) \\ &\geq 2\frac{1-\gamma}{\gamma}\log(1 + \epsilon_H) + p_H + \gamma s_H \end{aligned}$$

Now, as $2\log(1 + \epsilon_H) \geq \gamma s_H$ (since $2\log(1 + \epsilon_H) \geq p_H + \gamma s_H$ and $p_H \geq 0$), then the adversary does not reject the higher contract as $2\frac{1-\gamma}{\gamma}\log(1 + \epsilon_H) + p_H + \gamma s_H \geq p_H + s_H$. This also implies

$$(2 + 2(1-\gamma)/\gamma)\log(1 + \epsilon_H) \geq p_H + s_H \tag{2}$$

Then, the adversary either chooses the lower or higher contract. Then, the seller's utility is $(1-\rho)(q * (p_L + \gamma s_L) + (1-q) * (p_H + \gamma s_H)) + \rho[p_Z + s_Z - K(\exp(\epsilon_Z) - 1)]$, where Z is either L or H. If $Z = H$, then the revenue is $(1-\rho)(q * (p_L + \gamma s_L) + (1-q) * (p_H + \gamma s_H)) + \rho[p_H + s_H - K(\exp(\epsilon_H) - 1)]$. Then, observe that if $Z = L$, it means the adversary found L more attractive, that is, $-[p_L + s_L - K(\exp(\epsilon_L) - 1)] \geq -[p_H + s_H - K(\exp(\epsilon_H) - 1)]$. Thus, it can be said that $(1-\rho)(q * (p_L + \gamma s_L) + (1-q) * (p_H + \gamma s_H)) + \rho[p_H + s_H - K(\exp(\epsilon_H) - 1)]$ is an upper bound on the revenue.

Next, $(1-\rho)(q*(p_L + \gamma s_L) + (1-q)*(p_H + \gamma s_H)) + \rho(p_H + s_H)$ must be less than $p_L + \gamma s_L + p_H + s_H$, which by previous inequality number 2 and $p_L + \gamma s_L \leq \log(1 + \epsilon_H)$ is

$$\leq \quad \log(1 + \epsilon_H) + (2 + 2(1-\gamma)/\gamma)\log(1 + \epsilon_H) = 3\log(1 + \epsilon_H) + 2\frac{1-\gamma}{\gamma}\log(1 + \epsilon_H)$$

Now, choose $K = 10/\rho + 2\frac{1-\gamma}{\rho\gamma}$ which is clearly $\geq 2 + 2(1-\gamma)/\gamma$ also. Then,

$$\rho K[\exp(\epsilon_H) - 1] = (10 + 2(1-\gamma)/\gamma) * (\exp(\epsilon_H) - 1) \geq 10\log(1 + \epsilon_H) + 2\frac{1-\gamma}{\gamma}\log(1 + \epsilon_H)$$

Hence, the upper bound on revenue $(1-\rho)(q * (p_L + \gamma s_L) + (1-q) * (p_H + \gamma s_H)) + \rho[p_H + s_H] - \rho K(\exp(\epsilon_H) - 1)]$ is less than $-7\log(1 + \epsilon_H)$ which is strictly negative for positive ϵ_H, and thus, the revenue is negative. This contradicts the optimality of ϵ_H as 0 revenue is obtained with $p_H, \epsilon_H, s_H = 0$. We conclude that $\epsilon_H = \epsilon_L = 0$, so that $R^*_A = 0$.

On the other hand, without adversarial types, the operator can attain positive revenue. This is because the seller's problem (using Theorem 4.2) will be to maximize

$$q b_L(\epsilon_L) + (1-q)[b_H(\epsilon_H) - b_H(\epsilon_L) + b_L(\epsilon_L)] = b_L(\epsilon_L) - (1-q)b_H(\epsilon_L) + (1-q)b_H(\epsilon_H)$$

Hence, it is optimal to choose $\epsilon_H = 1$, leading to a lower bound of $R^* \geq (1-q)2\log 2$ on the non-adversarial revenue. □

Proof of Lemma 5.2. From Theorem 4.2, we know that $b_j(\epsilon_j) - p_j = b_j(\epsilon_{j-1}) - p_{j-1}$, or equivalently, $b_j(\epsilon_j) - b_j(\epsilon_{j-1}) = p_j - p_{j-1}$. Using the ID property of the benefit functions, for $i > j$ we get $b_i(\epsilon_j) - b_i(\epsilon_{j-1}) \geq b_j(\epsilon_j) - b_j(\epsilon_{j-1}) = p_j - p_{j-1}$, hence $b_i(\epsilon_j) - p_j \geq b_i(\epsilon_{j-1}) - p_{j-1}$. Thus, i prefers contract j to $j - 1$. Arguing inductively, we have the required result. □