Data-driven Contract Design

Parv Venkitasubramaniam and Vijay Gupta

Abstract— **We consider a game in which one player (the principal) seeks to incentivize another player (the agent) to exert effort that is costly to the agent. Any effort exerted leads to an outcome that is a stochastic function of the effort. The amount of effort exerted by the agent is private information for the agent and the principal observes only the outcome; thus, the agent can misreport his effort to gain higher payment. Further, the cost function of the agent is also unknown to the principal and the agent can also misreport a higher cost function to gain higher payment for the same effort. We pose the problem as one of contract design when both adverse selection and moral hazard are present. We show that if the principal and agent interact only finitely many times, it is always possible for the agent to lie due to the asymmetric information pattern and claim a higher payment than if he were unable to lie. However, if the principal and agent interact infinitely many times, then the principal can utilize the observed outcomes to update the contract in a manner that reveals the private cost function of the agent and hence leads to the agent not being able to derive any rent. The result can also be interpreted as saying that the agent is unable to keep his information private if he interacts with the principal sufficiently often.**

I. Introduction

Many problems in smart infrastructure networks can be expressed in the following form. A system operator wishes to incentivize one or multiple users to expend costly effort in order to take an action that is aligned with the operator (or the system) performance, but may not optimize the utility of the user directly. Each user is associated with an intrinsic parameter that determines both the cost that he incurs to expend a given effort and the quality of the output that results from a given effort that he expends. Since the intrinsic parameter as well as the amount of effort expended are unknown to the operator, the user may try to expend the least effort possible and falsify any reports of his parameter upon being asked. Designing a suitable incentive scheme for such agents has been considered in the literature with various assumptions and motivated by different application systems (see, e.g., [1]–[6] and the references therein).

Among the many possible applications to motivate the problem, we point out two. The first is that of the crowd-sensing or participatory sensing (see, e.g., [1]) in which an operator employs multiple sensors to generate and transmit measurements about an unknown value and uses these measurements to estimate the unknown value. The sensors expend energy or time to take these measurements. However, sensors may not directly benefit from ensuring that the operator generates an accurate estimate. To compensate the sensors for their effort cost, the sensors must be rewarded based on the accuracy of the information they provide. However, both the accuracy of the measurement and the effort cost are usually an increasing function of the intrinsic quality of the sensor, and both this intrinsic quality as well as the level of effort exerted by the sensors are private information for the sensors. Thus, the sensors have an incentive to expend little effort, yet misreport their effort and intrinsic quality in order to receive higher compensation [7].

The second application is that of demand response, in which an aggregator incentivizes customers to curtail their power usage in response to high peaks in demand or increase in transmission congestion increases [8]–[11]. The Federal Energy Regulatory Commission (FERC) defines demand response as the change in electric usage by end-use customers from their normal consumption patterns in response to changes in the price of electricity or any other incentive [12]. The effort put in by each customer puts to reduce her load is costly since it causes her discomfort. Each customer may be associated with an internal flexibility that characterizes the effort cost. Further, the amount of effort expended is private knowledge for each customer and the load reduction is only a noisy reflection of this amount of effort due to the effect of variables such as external temperature, actions of other users, and intrinsic characteristics of the customer himself. To incentivize the customer to put in ample effort even though it is costly, the aggregator must pay each customer proportional to the effort that the customer puts in. However, each customer can misreport his cost function and the amount of effort put in to gain more financial reward for the same load reduction [13]–[17].

Intuitively, designing an appropriate contract in this framework is difficult due to two reasons: (i) the agents do not benefit directly from taking the actions that the operator desires them to take, and (ii) the operator does not have access to either the intrinsic parameters or the amount of effort expended by the agent. The operator needs to design incentive mechanisms that mitigate both *moral hazard* (i.e., incentivizing desired actions by the agents when effort is costly and not observed by the principal, see, e.g., [18, Chapter 4]) and *adverse selection* (i.e., incentivizing agents to report their intrinsic parameters or 'types' truthfully, see, e.g., [18, Chapter 3]). While an extensive literature in contract theory (see, e.g., [1], [18], [19] and the references therein for an overview of the subject) has focused on resolving either moral hazard or adverse selection separately, or even moral hazard followed by adverse selection [20]–

This work is partly supported by the National Science Foundation through Awards CCF-1617889, CNS-1544724, ECCS-1550016, and CNS-1739295

Parv Vekitasubramaniam is with the Department of Electrical Engineering, Lehigh University, PA, USA pav309@lehigh.edu

Vijay Gupta is with the Department of Electrical Engineering, University of Notre Dam, IN, USA vgupta2@nd.edu

[23], the problem we consider features adverse selection followed by moral hazard. This problem has received much less attention in the literature with the few existing examples [18, Chapter 7.2] or [24] considering a *static* setting in which the principal and the agent interact only once. In a static setting, given that the operator does not have additional measurements and must rely solely on the information transmitted by the sensor, he must rely on some form of verification of the outcomes generated by the agents to constrain the agents not to expend the least possible effort and gain compensation corresponding to the highest effort cost function. It is known that verification can mitigate the information asymmetries of moral hazard and adverse selection (see e.g., [25]–[27], in which verification is similarly proposed) and many of the recent works have considered the problem with noisy, delayed, or infrequent verification [20]–[23], [28].

As against most of these works, we consider the problem when the principal and the agents interact repeatedly. Thus, our problem features adverse selection followed by moral hazard in a repeated setting. The repeated setting introduces new challenges since agents may adopt time-varying strategies and the operator may introduce a time-varying contract. If verification of the agent type or effort is done at every time step, one can devise a reputation-based scheme in which agents are paid based on both current outputs and reputations that summarize past interactions with the operator. Indeed, using reputation for mitigating information asymmetry, particularly in repeated games, is a popular strategy in the literature, see e.g. [29]–[35]. However, we do not assume verification of the agent type or effort at every time step. Instead, we use a data-driven approach in which we use only the history of the observed outputs to update the contract offered by the operator. With only the observed outputs, it is not immediate if the operator can even limit, let alone eliminate, the incentive for the agent to either lie or not put in the desired effort. Perhaps somewhat surprisingly, we show that in the limit of infinitely many interactions, the operator can indeed do so. For a finite number of interactions, we characterize the rent that an agent can derive from the absence of verification. The basic insight on which the proofs are based is that the history of outputs provides knowledge of the 'quality' of the work that an agent does, and hence limits the amount of falsification he can introduce. However, since the agent can always degrade his effort, care must be taken to use this knowledge in identifying the type or the parameters of the agent correctly. The result can also be interpreted as a privacy result in that it says that private information from an agent will always be revealed to the operator in the limit of infinitely many interactions (unless the agent is willing to suffer an infinite amount of loss).

The paper is organized as follows. The mathematical formulation of the leader-follower game is described in Section II. The one-shot version of the game is studied in Section III . The dynamic game is studied in Section IV with particular emphasis on the asymptotic behavior. A special case of the problem, when quality of work is not impacted by adverse selection, is showcased in Section V where it is shown that rent cannot be eliminated entirely. Finally, some concluding remarks are presented in Section VI.

II. Mathematical Formulation

We consider a repeated leader follower game played between a risk-neutral principal and a risk-neutral agent whose attributes are described in the following.

Principal The principal is the leader whose action is to issue a contract defined by a function $w : \mathbb{R} \mapsto \mathbb{R}$, where $w(x)$ is the payment offered to the agent for his effort when the observable output is given by $x \in \mathbb{R}$. In a single instance of the game, the principal's utility, should the agent accept the contract, is measured by a function $S : \mathbb{R} \mapsto \mathbb{R}$ where $S(x)$ is the utility from the observable output. The payoff for the principal is then given by $S(x) - w(x)$ should the contract be accepted by the agent, and 0 if rejected.

Agent The agent is the follower who responds to the issued contract, by accepting or rejecting it, and if accepted, takes an effort $a \in \mathcal{A}$, a compact subset of $\mathbb{R}^+$. Prior to the start of the repeated game, the agent obtains private information defined by a type variable $\theta \in \Theta$ which influences his utility and the produced output. More specifically, the output X is a random variable given by the function $q_X : \mathcal{A} \times \Theta \times \mathbb{R} \mapsto \mathbb{R}$, where $q_X(a, \theta, N)$ depends on, aside from the effort and type, a random noise variable N distributed according to pdf q_N. The payoff for the agent, should he accept the contract, is given by $w(x) - h(a, \theta)$ where $h : \mathcal{A} \times \Theta \mapsto \mathbb{R}$ is the cost incurred as a function of the agent's type and effort.

Game Formulation The game is played over several time steps, and our work primarily looks into the equilibria as the time horizon approaches infinity. The format of the multi-period game is described below:

1) $t = 0$: Agent obtains value of θ. Let $k = 0$.
2) $t = 3k + 1$: Principal issues a contract defined by a function $w_k(\cdot)$.
3) $t = 3k + 2$: The agent chooses to accept or reject the contract. If he accepts, exerts an effort a_k which could be the realization of a random variable A_k.
4) $t = 3k + 3$: If contract $w_k(\cdot)$ was accepted, the effort A_k results in an output $X_k = q_X(A_k, \theta, N_k)$ where the random noise variables $\{N_k\}$ are independent and identically distributed according to q_N. Based on the output, the principal and agent receive their respective payoffs $U_p^k = S(X_k) - w(X_k)$ and $U_a^k = w(X_k) - h(A_k, \theta)$. If the contract were rejected, both parties receive zero payoff.
5) $k = k + 1$ and repeat steps 2-4 if $k < K$.

K is the horizon of the game. The case of $K = 1$ is the single step game. K is finite for a finite horizon game and $\to \infty$ for an infinite horizon game.

The principal is not privy to the agent's private information θ except through a prior distribution p_θ over Θ, which for the purposes of this work is assumed to be finite. The objective of the principal is to design a sequence of contracts $w_1(\cdot), w_2(\cdot), \cdots$ to maximize the expected cumulative payoff

$$U_p = \frac{\mathbb{E}\left(\sum_k U_p^k\right)}{K} \tag{1}$$

where the expectation is over the noise, action and type variables. For an agent of type θ, the objective is to choose acceptance and rejection of each contract, and the consequent action A_k upon acceptance, such that the conditional expected cumulative payoff is maximized:

$$U_a = \frac{\mathbb{E}\left(\sum_k U_a^k|\theta\right)}{K} \tag{2}$$

In this work, we study the achievable equilibria in this problem, such that both principal and agent satisfy their individual rationality conditions. In the subsequent sections, we investigate different facets of this problem by making specific assumptions on the output, cost and utility functions as follows:

- $S(x) = x$
- $h(a,\theta) = \frac{a^2}{2\theta}$
- $X_k = A_k + Q(\theta)N_k$ where $N_k \sim \mathcal{N}(0,\sigma^2)$ and $Q(\cdot)$ is an increasing function

These assumptions, while specific to enable explicit characterization of equilibria, are representative of key qualitative aspects of the problem namely, the principal's utility is an increasing function of the produced output, the agent's cost is an increasing function of the effort, and a decreasing function of the type variable which inversely relates to quality. Consequently, higher the value of θ, higher the noise variance exemplifying a lower quality of produced output.

Under these assumptions, in Section III we derive the equilibrium strategies for the players and show that in the single step game, some agents can extract rent (gain additional reward by exploiting the principal's incomplete information) and it is impossible for the principal to eliminate the loss due to the adverse selection problem. Although we present the result and the proof for a single step game, it can readily be extended to any finite horizon game. In Section IV, we study the infinite horizon version of the game, and demonstrate that as long as the produced output is impacted by the agent's type, the principal can in the long run learn the type through observable outputs – a noisy verification process – and incentivize the agent to exert the desired effort at asymptotically negligible rent. If, however, the quality of work is independent of type, in other words $Q(\cdot)$ is independent of θ, it is possible for an agent of a higher type to impersonate one of a lower type without being detected and continue to extract infinite rent asymptotically.

III. One Shot Contract

The one shot game between the principal and agent can be formulated using standard individual rationality and incentive compatibility constraints as:

$$\mathcal{P}: \begin{cases} & \max_{w(\cdot)} \mathbb{E}(S(X) - w(X)) \\ \text{s.t.} & \text{IC: } \mathbb{E}(w(X)|\theta) - h(a^*,\theta) \geq 0 \\ \text{where} & \text{IR: } a^* = \arg\max_a \mathbb{E}(w(X)|\theta) - h(a^*,\theta) \end{cases} \tag{3}$$

Since Θ is finite, without loss of generality we assume $\Theta = \{\theta_1, \theta_2, \cdots, \theta_M\}$ where $\theta_1 \leq \theta_2 \cdots \leq \theta_M$.

Contracts in general can take any functional form that depends on the observed output. We limit ourselves to a standard linear form $w(x) = x - d$ where d is a constant. The goal of the principal is to then characterize the optimal value of d that solves (3). Were the principal to be aware of the type θ of the agent (i.e. no adverse selection), then it is easily shown that the optimal contract:

$$w(x) = x - \frac{\theta}{2}$$

where the agent's IC condition is satisfied and the optimal effort $a^* = \theta$ which is also equal to the first best effort from the principal's perspective thus allowing zero rent. Under adverse selection however, it is not always possible to completely avoid rent, as is shown in the following theorem.

Theorem 1: In the one shot game, the optimal linear contract offered by the principal is given by:

$$w(x) = x - \max_m \frac{\theta_m}{2} \sum_{i=m}^{M} p_\theta(\theta_i) \tag{4}$$

and agents whose type θ_i satisfies $i \geq \arg\max_m \theta_m \sum_{i=m}^{M} p_\theta(\theta_i)$ choose to accept the contract and exert an effort $a = \theta_i$

Proof: The proof is an application of standard first order conditions and is available in the Appendix. □

Under the optimal contract, amongst the types of agents who accept the contract, those whose type θ_i satisfy $i > \arg\max_m \theta_m \sum_{i=m}^{M} p_\theta(\theta_i)$ are able to extract rent. The loss L_1 from the principal's perspective due to the adverse selection and moral hazard in the one-shot game is then:

$$L_1 = \frac{\mathbb{E}(\theta) - \max_m \theta_m \sum_{i=m}^{M} p_\theta(\theta_i)}{2}$$

Note that the above argument can be generalized to any finite horizon game at the expense of more notation. In the subsequent section we consider the infinite horizon game and present the argument that the loss $\frac{L_K}{K}$ for K repeated games goes to zero as $K \to \infty$ as long as the agent's type impacts the quality of output.

IV. Dynamic Contract

We now focus our attention to the infinite horizon framework and build an argument that continually observed outputs allows the principal to learn about the agent's private information and consequently offer the contract that extracts the first best effort at zero rent. We will consider two scenarios to illustrate the impact of noisy verification; first when the quality of work factor $Q(\theta)$ is an increasing one-one function of θ, and second when $Q(\theta) = 1, \forall\theta$. In the former, since the agent's type impacts the observed variance in output, the principal can use a hypothesis test to detect (asymptotically accurately) the type, and eliminate the rent by offering the best contract for the agent of the detected type. In the latter, since verification is impossible, we will show in the subsequent section that it is possible for agents to fake their types and secure infinite rent across the horizon.

Consider the dynamic contracting problem when $Q(\theta) = \theta$, in other words, the produced output for contract $w_k(\cdot)$ is:

$$X_k = A_k + \theta N_k$$

where $N_k \sim \mathcal{N}(0, \sigma^2)$. This is not a limiting assumption; were $Q(\theta)$ to be defined as any increasing one-one function, the proofs can be easily modified, and the results would hold. We divide the principal's strategy into two phases, a detection phase, and an execution phase. During the detection phase, the principal uses the following hypothesis test on a sequence of n observations of the produced outputs $X_1, \cdots, X_n$:

$$\delta(X_1, \cdots, X_n) = \begin{cases} H_m & m = \min\left\{i : \left|\bar{X} - \theta_i\right| < \epsilon \right. \\ & \left. \text{and } \left|\hat{\sigma_X}^2 - \theta_i^2 \sigma^2\right| < \epsilon\right\} \\ H_M & \nexists i : \left|\bar{X} - \theta_i\right| < \epsilon \\ & \text{and } \left|\hat{\sigma_X}^2 - \theta_i^2 \sigma^2\right| < \epsilon \end{cases} \tag{5}$$

where

$$\bar{X} = \frac{\sum_{k=1}^{n} X_k}{n}, \hat{\sigma_X}^2 = \frac{\sum_{k=1}^{n} (X_k - \bar{X})^2}{n-1}$$

and H_m refers to the hypothesis that the agent's type is θ_m. During the detection phase $k \leq n$, the principal offers the contract

$$w_k(x) = x - \frac{\theta_1}{2}$$

to incentivize participation from every type of agent. Following the detection phase, based on the observed outputs, the principal uses the above rule to pick a hypothesis, let's say H_m, and for the rest of the horizon $n < t \leq K$ offers a contract

$$w_k = x - \frac{\theta_m}{2}$$

The **two-phase dynamic contract** strategy is also detailed in Algorithm 1. In the following theorem we show that this strategy enables the principal to extract the first best effort from every type of agent at asymptotically negligible rent.

Theorem 2: Under the two-phase dynamic contract strategy specified in Algorithm 1, if $Q(\theta)$ is a one-one function of θ, the average cumulative payoff for the principal over K repeated games, as $K \to \infty$ is given by:

$$U_P = \lim_{K \to \infty} \frac{\sum_{k=1}^{K} \mathbb{E}(U_P^k)}{K} = \frac{\mathbb{E}(\theta)}{2}$$

Proof: The proof relies on the fact that as K grows larger, as long as the length of the detection phase $n \in o(K)$, the statistical measures of the observed outputs should concentrate around the true measures of the underlying distribution which are sufficiently separated for different values of θ. Since the default hypothesis is the "worst" contract, it disincentivizes agents to fake their type. The mathematical details of the proof are available in the Appendix. □

Consequent to Theorem 2, we can state that in the long run, the ability of the principal to learn private information implicitly through the observed outputs allows her to eliminate the effect of adverse selection and thus extract first best efforts at zero rent. The observed outputs in this scenario play the role of a noisy verification channel which in the long run is increasingly accurate. In the subsequent section, we underline this idea further by considering an output process independent of type, and show that in that setup, it is possible for agents to fake their types forever.

Algorithm 1 Two Phase Dynamic Contracting Strategy

$k \leftarrow 1$
for $k < n$ **do**
  Principal issues contract $w_k(x) = x - \frac{\theta_1}{2}$
  if Agent accepts contract and performs action A_k **then**
    $X_k \leftarrow A_k + \theta N_k$ where θ is type of agent
  end if
  $k \leftarrow k + 1$
end for
if Agent rejected any contract $w_k(\cdot)$ for $k < n$ **then**
  $\hat{\theta} = \theta_M$
else
  $H_m \leftarrow \delta(X_1, \cdots, X_n)$
  $\hat{\theta} = \theta_m$
end if
for $k > n$ **do**
  Principal issues contract $w_k(x) = x - \frac{\hat{\theta}}{2}$
  if Agent accepts contract and performs action A_k **then**
    $X_k \leftarrow A_k + \theta N_t$ where θ is true type of agent
  end if
  $k \leftarrow k + 1$
end for

V. Rent through Pretense

In this section, we limit ourselves to binary private information $\Theta = \{\theta_{low}, \theta_{high}\}$ such that $\theta_{low} < \theta_{high}$, and let $p_\theta(\theta_{low}) = p$. From Theorem 1, we know that the optimal contract in the one-shot game is given by:

$$w(x) = \begin{cases} x - \frac{\theta_{low}}{2} & \theta_{low} > \theta_{high}(1-p) \\ x - \frac{\theta_{high}}{2} & \text{o.w} \end{cases} \tag{6}$$

In this one shot game, when the condition $\theta_{low} > \theta_{high}(1-p)$ is satisfied (and consequently both types of agents accept the contract), the principal's payoff is $\frac{\theta_{low}}{2}$ and the θ_{high} agent is able to extract rent equal to $\frac{\theta_{high} - \theta_{low}}{2}$. The following theorem will demonstrate that, when quality of output is not impacted by private information ie $Q(\theta) = 1$, the rent does not completely vanish even in the infinite horizon case by virtue of the θ_{high} agent pretending to have a type θ_{low} without ever being detected.

Theorem 3: In a repeated leader-follower game between a principal and agents with binary types $\Theta = \{\theta_{low}, \theta_{high}\}$, if $Q(\theta) = 1$, the equilibrium average cumulative payoff for the principal is given by:

$$U_P = \lim_{K \to \infty} \frac{\sum_{k=1}^{K} \mathbb{E}(U_P^k)}{K} = \frac{\theta_{low}}{2} \tag{7}$$

and the optimal strategy for both types of agents is to exert effort $A_k = \theta_{low}$ in every time step.

Proof: It is easy to see that as long as both type of agents adopt identical strategies, the observed outputs are statistically indistinguishable between the two types of agents. Since the agent of type θ_{low} has no incentive to alter his one-shot strategy, the only was the agent of type θ_{high} obtains rent is by faking his type. It only remains to be seen what the optimal contractual strategy is for the principal that results in an equilibirium with these agent strategies. The key idea is to use a repeated hypothesis test to ensure that the agents do not deviate from the strategy $A_k = \theta_l$. Further details are available in the appendix. □

Consequent to Theorem 3, the principal's expected payoff per contract is unchanged between the single step and infinite horizon frameworks. The payoff of the agent of type θ_{high} per contract is slightly reduced (from $\frac{\theta_{high}-\theta_{low}}{2}$ to $\frac{\theta_{low}}{2}\left(1-\frac{\theta_{low}}{\theta_{high}}\right)$) due to the need for pretense. In other words, even though the availability of data does not improve the per contract payoff of the principal it does reduce the rent extracted. We do state a caveat here that the fact that long term observations does not improve the principal's payoff is a consequence of the linear payment form which results in a payoff independent of output. Were the payment formulated differently, we believe it would improve the payoff even though the rent would not be eliminated entirely as long as pretense is profitable to the agent.

Theorems 1 and 2 together suggest a tradeoff between privacy and efficiency, albeit using the two extreme scenarios (one-shot and infinite horizon). In general, for a finite horizon model, complete analytical characterization of the equilibria is intractable even for the simplest assumptions on parameters. In Figure 1, we plot the agent's reward as a function of the horizon using numerical stochastic optimization (assuming $\Theta = \{1, 2\}, \sigma = 1$) by limiting ourselves to **deterministic** agent policies. The effect of data on the efficiency (rent reduction), in other words the privacy-efficiency tradeoff, and the importance of the quality factor, are visibly demonstrated in the plot.

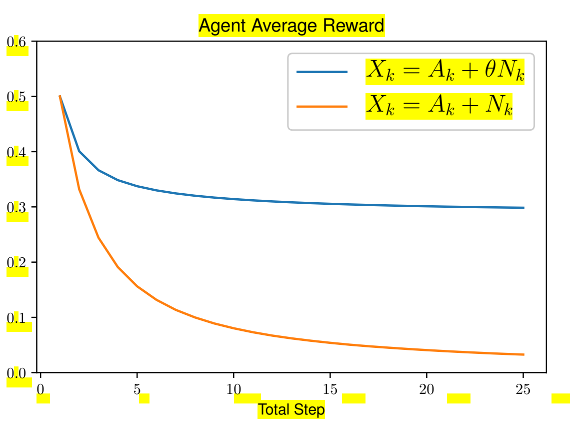

Fig. 1. Rent extracted vs Horizon

VI. Concluding Remarks

In this paper, we considered the problem of contract design when a principal and an agent interact repeatedly in the presence of adverse selection followed by moral hazard. We show that there is a fundamental difference in situations when the intrinsic type of the agent also determines the quality of the output of the effort and when they do not, as also when the game is played over a finite horizon versus an infinite horizon. Specifically, we show that the principal must pay a rent to derive the first best outcome if the quality of the output of the effort is independent of the type or if the game is played over a finite horizon. On the other hand, if the quality of the output depends on the type of the agent and if the agent and principal interact infinitely often, then any private information of the agent must be revealed to the principal and the agent cannot derive any rent.

The work can be extended along multiple lines. One direction is to apply these results to specific examples by considering any additional constraints imposed by the application. Another direction is to consider the situation when multiple agents are present, who may collude to hide information from the principal. For instance, a high quality agent can pretend to be low quality so that the principal cannot differentiate among the agents and must pay rent. It will be interesting to characterize the rent that such coalitions can produce.

References

[1] H. Gao, C. H. Liu, W. Wang, J. Zhao, Z. Song, X. Su, J. Crowcroft, and K. K. Leung, "A survey of incentive mechanisms for participatory sensing," *Communications Surveys & Tutorials, IEEE*, vol. 17, no. 2, pp. 918–943, 2015.

[2] Y. Zhang and M. van der Schaar, "Collective ratings for online communities with strategic users.," *IEEE Trans. Signal Processing*, vol. 62, no. 12, pp. 3069–3083, 2014.

[3] W. Saad, Z. Han, H. V. Poor, and T. Basar, "Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 86–105, 2012.

[4] C. G. Cassandras and W. Li, "Sensor networks and cooperative control," *European Journal of Control*, vol. 11, no. 4, pp. 436–463, 2005.

[5] R. Borndörfer, J. Buwaya, G. Sagnol, and E. Swarat, "Optimizing toll enforcement in transportation networks: a game-theoretic approach," *Electronic Notes in Discrete Mathematics*, vol. 41, pp. 253–260, 2013.

[6] M. H. Manshaei, Q. Zhu, T. Alpcan, T. Bacşar, and J.-P. Hubaux, "Game theory meets network security and privacy," *ACM Computing Surveys (CSUR)*, vol. 45, no. 3, p. 25, 2013.

[7] D. G. Dobakhshari, P. Naghizadeh, M. Liu, and V. Gupta, "A reputation-based contract for repeated crowdsensing with costly verification," in *American Control Conference (ACC), 2017*, pp. 5243–5248, IEEE, 2017.

[8] A. C. Tellidou and A. G. Bakirtzis, "Agent-based analysis of capacity withholding and tacit collusion in electricity markets," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1735–1742, 2007.

[9] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, "A survey on demand response programs in smart grids: Pricing methods and optimization algorithms," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 152–178, 2015.

[10] R. Deng, Z. Yang, M.-Y. Chow, and J. Chen, "A survey on demand response in smart grid: Mathematical models and approaches," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 3, pp. 570–582, 2015.

[11] M. H. Albadi and E. F. El-Saadany, "Demand response in electricity markets: An overview," in *Power Engineering Society General Meeting, 2007. IEEE*, pp. 1–5, IEEE, 2007.

[12] V. M. Balijepalli, V. Pradhan, S. Khaparde, and R. Shereef, "Review of demand response under smart grid paradigm," in *Innovative Smart Grid Technologies-India (ISGT India), 2011 IEEE PES*, pp. 236–243, IEEE, 2011.

[13] H.-p. Chao, "Price-responsive demand management for a smart grid world," *The Electricity Journal*, vol. 23, no. 1, pp. 7–20, 2010.

[14] H.-p. Chao and M. DePillis, "Incentive effects of paying demand response in wholesale electricity markets," *Journal of Regulatory Economics*, vol. 43, no. 3, pp. 265–283, 2013.

[15] Y. Chen, W. S. Lin, F. Han, Y.-H. Yang, Z. Safar, and K. R. Liu, "A cheat-proof game theoretic demand response scheme for smart grids," in *Communications (ICC), 2012 IEEE International Conference on*, pp. 3362–3366, IEEE, 2012.
[16] D. G. Dobakhshari and V. Gupta, "A contract design approach for phantom demand response," *IEEE Transactions on Automatic Control*, 2018.
[17] D. Muthirayan, D. Kalathil, K. Poolla, and P. Varaiya, "Mechanism design for self-reporting baselines in demand response," in *American Control Conference (ACC), 2016*, pp. 1446–1451, IEEE, 2016.
[18] J.-J. Laffont and D. Martimort, *The theory of incentives: the principal-agent model*. Princeton university press, 2009.
[19] B. Salanié, *The economics of contracts: a primer*. MIT press, 2005.
[20] A. Dasgupta and A. Ghosh, "Crowdsourced judgement elicitation with endogenous proficiency," in *Proceedings of the 22nd international conference on World Wide Web*, pp. 319–330, ACM, 2013.
[21] J. A. Bohren and T. Kravitz, "Optimal contracting with costly state verification, with an application to crowdsourcing," 2016. Working paper available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2889031.
[22] K. J. Crocker and T. Gresik, "Optimal compensation with earnings manipulation: Managerial ownership and retention," tech. rep., mimeo, 2010. mimeo, https://www3.nd.edu/∼tgresik/Compensation.pdf.
[23] D. G. Dobakhshari, N. Li, and V. Gupta, "An incentive-based approach to distributed estimation with strategic sensors," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pp. 6141–6146, IEEE, 2016.
[24] J. Poblete and D. Spulber, "The form of incentive contracts: agency with moral hazard, risk neutrality, and limited liability," *The RAND Journal of Economics*, vol. 43, no. 2, pp. 215–234, 2012.
[25] D. Mookherjee and I. Png, "Optimal auditing, insurance, and redistribution," *The Quarterly Journal of Economics*, vol. 104, no. 2, pp. 399–415, 1989.
[26] K. C. Border and J. Sobel, "Samurai accountant: A theory of auditing and plunder," *The Review of economic studies*, vol. 54, no. 4, pp. 525–540, 1987.
[27] D. P. Baron and D. Besanko, "Regulation, asymmetric information, and auditing," *The RAND Journal of Economics*, pp. 447–470, 1984.
[28] D. G. Dobakhshari, P. Naghizadeh, M. Liu, and V. Gupta, "A reputation-based contract for repeated crowdsensing with costly verification," 2017. Extended Version, available at https://www3.nd.edu/∼dghavide/ACC7.pdf.
[29] P. Naghizadeh and M. Liu, "Perceptions and truth: a mechanism design approach to crowd-sourcing reputation," *IEEE/ACM Transactions on Networking (TON)*, vol. 24, no. 1, pp. 163–176, 2016.
[30] G. J. Mailath and L. Samuelson, *Repeated games and reputations: long-run relationships*. Oxford university press, 2006.
[31] N. Cao, S. Brahma, and P. K. Varshney, "Target tracking via crowdsourcing: A mechanism design approach.," *IEEE Trans. Signal Processing*, vol. 63, no. 6, pp. 1464–1476, 2015.
[32] D. Abreu, D. Pearce, and E. Stacchetti, "Toward a theory of discounted repeated games with imperfect monitoring," *Econometrica: Journal of the Econometric Society*, pp. 1041–1063, 1990.
[33] A. Bohren, "Stochastic games in continuous time: Persistent actions in long-run relationships," tech. rep., working paper, 2012.
[34] C. Dellarocas, "Reputation mechanism design in online trading environments with pure moral hazard," *Information Systems Research*, vol. 16, no. 2, pp. 209–230, 2005.
[35] K. Han, H. Huang, and J. Luo, "Posted pricing for robust crowdsensing," in *Proceedings of the 17th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pp. 261–270, ACM, 2016.

Appendix

Proof of Theorem 1

Due to individual rationality of the agent and the specific linear form of the contract, if an agent accepts the contract, then he will take an action a^* that satisfies:

$$a^* = \arg\max \mathbb{E}\left(X - d - \frac{a^2}{2\theta}|\theta\right)$$

Since $\mathbb{E}(X) = a$, the optimal action for the agent is given by $a^* = \theta$. The incentive compatibility criterion then reduces to thecondition $\frac{\theta}{2} > d$ and the principal's optimization problem takes the form:

$$\max_d d \Pr\{\frac{\theta}{2} > d\}$$

which for a finite Θ reduces to the expression in (4). □

Proof of Theorem 2

Since $Q(\theta)$ is an increasing one-one function independent of the time horizon n, without loss of generality we can assume that $Q(\theta) = \theta$. If the detection phase results in accurately identifying the type of the agent, then the contract issued by the principal in the execution phase secures first best effort from the agent at zero rent. It remains to be seen how the probability of incorrect detection falls as the horizon $K \to \infty$. The key to the proof is bounding the probabilities of the following events for the true and any alternate hypotheses:

$$E_{1m} : \left|\frac{\sum X_k}{n} - \theta_m\right| < \epsilon\}$$
$$E_{2m} : \left|\frac{\sum (X_k - \bar{X})^2}{n-1} - \theta_m^2\sigma^2\right| < \epsilon$$

More specifically, we will show that if the true type $\theta = \theta_m$, then $\Pr\{E_{1m}^c \bigcup E_{2m}^c\} \leq \frac{\kappa_1}{n}$ (missed detection) for some constant κ_1, and if $\theta > \theta_m$, then $\Pr\{E_{1m} \bigcup E_{2m}\} \leq \frac{\kappa_1}{n}$ (false alarm).

When the true type $\theta = \theta_m$, applying standard Chebyshev inequalities on the Gaussian random variable $\frac{\sum X_k}{n}$ and the Chi-squared variable $\frac{\sum (X_k - \bar{X})^2}{n-1}$ we can bound the missed detection probability as:

$$\begin{aligned} \Pr\{E_{1m}^c \bigcup E_{2m}^c\} &\leq \Pr\{E_{1m}^c\} + \Pr\{E_{2m}^c\} \\ &\leq \frac{\theta_m^2\sigma^2}{n} + \frac{2\theta_m^4\sigma^4}{\epsilon^2(n-1)} \leq \frac{\kappa_1}{n} \end{aligned} \quad (8)$$

for some constant κ_1.

When $\theta = \theta_i > \theta_m$, we will assume that condition E_{1m} is satisfied with probability 1, and bound the probability that E_{2m} is also true. Since false alarm in this case can only lead to a loss for the principal, the payoff computed under this assumption will serve as a lower bound for the actual expected payoff. Let $A_1, \cdots, A_n$ be the random variables denoting the sequence of actions taken by the θ_i agent in response to the received sequence of contracts in the learning phase. For large n, we bound $\Pr\{E_{2m}\}$ as

$$\begin{aligned} &\Pr\left\{\left|\frac{\sum (X_k - \bar{X})^2}{n-1} - \theta_m^2\sigma^2\right| < \epsilon\right\} \\ = &\Pr\left\{\left|\frac{\sum (X_k - \bar{X})^2}{n-1} - \theta_m^2\sigma^2\right| < \epsilon\right\} \\ = &\Pr\left\{\left|\frac{\sum (A_k + N_k - \bar{X})^2}{n-1} - \theta_m^2\sigma^2\right| < \epsilon\right\} \\ \leq &\Pr\left\{\left|\frac{\sum N_k}{n}\right| > \epsilon\right\} + \Pr\left\{\left|\frac{\sum A_k N_k}{n}\right| > \epsilon\right\} \\ &+ \Pr\left\{\left|\frac{\sum A_k^2 + N_k^2 - \bar{X}^2}{n-1} - \theta_m^2\sigma^2\right| < \epsilon\right\} \end{aligned}$$

Since N_k are iid zero mean Gaussian, for large n, the first two probabilities in the above equation fall exponentially with n. We use a one sided Chebyshev inequality to bound the third term. Let

$$\tilde{\sigma}_A^2 = \frac{\sum_k A_k^2 - \bar{A}^2}{n-1}, \; \tilde{\sigma}_N^2 = \frac{\sum_k N_k^2 - \bar{N}^2}{n-1}$$

Then for n large enough

$$\begin{aligned}
& \Pr\left\{\left|\frac{\sum A_k^2 + N_k^2 - \bar{X}^2}{n-1} - \theta_m^2\sigma^2\right| < \epsilon\right\} \\
\leq\; & \Pr\left\{\left|\tilde{\sigma}_N^2 + \tilde{\sigma}_A^2 - \sigma^2\theta_m^2\right| < 2\epsilon\right\} \\
\leq\; & \Pr\left\{\tilde{\sigma}_N^2 < \sigma^2\theta_m^2 + 2\epsilon\right\} \\
=\; & \Pr\left\{\frac{\sum_k (N_k - \bar{N})^2}{\sigma^2\theta_i(n-1)} < \frac{\theta_m^2}{\theta_i^2} + 2\frac{\epsilon}{\theta_i^2}\right\} \\
=\; & \Pr\left\{\frac{\sum_i (N_k - \bar{N})^2}{\sigma^2\theta_i^2} - (n-1) < (n-1)\left(\frac{2\epsilon}{\theta_i^2\sigma^2} - \left(1 - \frac{\theta_m^2}{\theta_i^2}\right)\right)\right\} \\
\leq\; & \frac{\kappa_2}{n}
\end{aligned} \tag{9}$$

for n large enough using the Chebyshev-Cantelli Theorem (one-sided Chebyshev) on the sample variance of an iid Gaussian sequence. In other words, regardless of the actions of the agent, the probability that he can pretend to be of type θ_m falls as $o(1/n)$.

Equations (8) and (9) guarantee that as $n \to \infty$, the probability of incorrect detection falls as zero. Therefore, as long the length of the detection phase n satisfies the conditions $n \in o(K)$ and $n \to \infty$ as $K \to \infty$, the loss incurred by the principal due to issuing an incorrect contract in the execution phase falls to 0 as the total horizon $K \to \infty$. □

Proof of Theorem 3

We only consider the scenario $\theta_{low} > \theta_{high}(1-p)$ when it is in the principal's best interest to incentivize both types of agents to accept the contract. Let K be the length of the horizon, we shall eventually let $K \to \infty$. The principal adopts the following strategy.

1) For $n \in o(K)$ time steps, the principal offers the contract $P(x) = x - \frac{\theta_{low}}{2}$, and observes the sequence of outputs.
2) Assuming that the actions taken by the agent were identical across steps and equal to the optimal action in the single step game, ie $A_k = \theta$, the principal performs a hypothesis test on the sequence of outputs to conclude if $\theta = \theta_{low}$ or θ_{high}.
3) If the hypothesis test resulted in the type being identified as θ_{low}, the principal continues to offer the contract $P(x) = x - \frac{\theta_{low}}{2}$ for a further n steps, else the principal offers the contract $P(x) = x - \frac{\theta_{high}}{2}$ to the agent for a further n steps.
4) Steps 2 and 3 are repeated till the end of the horizon K

Note that as $K \to \infty$, were the agents to stick to their optimal one-shot strategy repeatedly, $a = \theta$ through an $o(N)$ phase, the hypothesis test yields asymptotically a perfect identification of the agent's type.

For the agent of type θ_{low}, since only two possible contracts are on offer, it is in his best interest to stick to the action $A_k = \theta_{low}$ at all time steps.

For the agent of type θ_{high}, there are two possibilities, reveal his type by taking action $A_k = \theta_{high}$ forever, or pretend to be an agent of type θ_{low} by taking action $A_k = \theta_{low}$ forever. As it turns out, the honest strategy fetches zero rent for this agent, whereas the pretense strategy secures a limiting average rent of $\frac{\theta_{low}}{2}\left(1 - \frac{\theta_{low}}{\theta_{high}}\right)$ which would then be the best choice. It is easy to see that given the chosen strategies of the agents, the principal has no reason to change her strategy thus guaranteeing equilibrium. □