Robust Automated Mechanism Design[1]

MICHAEL ALBERT, University of Texas at Austin
VINCENT CONITZER, Duke University
PETER STONE, University of Texas at Austin

We introduce a new class of mechanisms, *robust mechanisms*, that is an intermediary between ex-post mechanisms and Bayesian mechanisms. This new class of mechanisms allows the mechanism designer to incorporate imprecise estimates of the distribution over bidder valuations in a way that provides strong guarantees that the mechanism will perform at least as well as ex-post mechanisms, while in many cases performing better. We further extend this class to mechanisms that are with high probability incentive compatible and individually rational, ϵ-*robust mechanisms*. Using techniques from automated mechanism design and robust optimization, we provide an algorithm polynomial in the number of bidder types to design robust and ϵ-robust mechanisms. We show that while no mechanism in this class can guarantee better performance than an ex-post mechanism in all cases, experimentally this new class of mechanisms can significantly outperform traditional mechanism design techniques when the mechanism designer has an estimate of the distribution and the bidder's valuation is correlated with an externally verifiable signal.

1. INTRODUCTION

Auctions are one of the fundamental tools of the modern economy for allocating resources. They are used to allocate online ad space, offshore oil drilling rights, famous artwork, small and medium lift capacity to planetary orbit, government supply contracts, FCC spectrum licenses, and almost limitless numbers of other things, large and small. Further, the sizes of these markets are economically enormous. In 2014, \$10 billion dollars of ad revenue was generated through automated auctions [Interactive Advertising Bureau (IAB) 2015]. In 2012, just four government agencies—the Army, the Department of Homeland Security, the Department of the Interior, and the Department of Veteran Affairs—purchased \$800+ million of commercial items through auctions [Government Accountability Office 2013]. In 2014, NASA awarded contracts to Boeing and Space-X worth \$4.2 billion and \$2.6 billion respectively through an implicit auction process [NASA 2014]. In an upcoming auction this year, the FCC is expected to allocate between \$60 and \$80 billion worth of broadcast spectrum. Given the economic magnitudes involved, it is crucial that these auctions are implemented optimally, for even small deviations from optimality can lead to millions of dollars worth of lost revenue, inefficiencies in resource allocation, and overspending.

It has long been understood that revenue optimal auction mechanisms are *prior-dependent* (also known as *Bayesian*) mechanisms [Myerson 1981; Cremer and McLean 1985, 1988; Lopomo 2001], i.e., mechanisms that assume some knowledge of the kinds of bidders that are likely to participate. For a seller trying to maximize revenue by selling a single item to independent bidders, she[2] needs to set a reserve price below which she will not sell the item [Myerson 1981], and this reserve price is dependent on her belief about the bidders she is likely to face. However, much of the focus of the algorithmic mechanism design community has been on the approximate optimality of *simple* mechanisms [Bulow and Klemperer 1996; Hartline and Roughgarden 2009; Roughgarden and Talgam-Cohen 2013; Morgenstern and Roughgarden 2015], that is, mechanisms that are either prior-independent or weakly prior-dependent (this distinction will be made clear later on). This focus is due to two factors. First, prior-dependent mechanisms can be very brittle to mis-specified priors [Lopomo 2001; Albert et al. 2015]. That is to say, if a prior-dependent mechanism is constructed using an incorrect

[1] In *EC 2016 Second Workshop on Algorithmic Game Theory and Data Science*

[2] In this paper, we will use "he" to denote bidders and "she" to denote mechanism designers/sellers.

prior it can perform much worse than simple mechanisms [Hartline and Roughgarden 2009]. Second, competition can be an effective substitute for knowledge of the distribution [Bulow and Klemperer 1996], so instead of implementing prior-dependent mechanisms, practitioners generally implement prior-independent mechanisms under the assumption that there are many bidders, or that it will somehow be possible to *acquire* new bidders

Unfortunately, in many auctions there is no feasible way to acquire more bidders. When NASA is awarding contracts to private space companies, they cannot generate more companies with the expertise to provide lift capacity, nor can Google find more bidders who are interested in advertising on search results for obscure brand-related keywords. Thin auctions are an actively recognized concern for many organizations that use auction mechanisms. A Government Office of Accountability report from 2013 [Government Accountability Office 2013] examining the use of reverse auctions by four governmental organizations, found that of the 19,688 reverse auctions the organizations conducted in 2012 with a total worth of \$800+ million, over one-third had only a *single bidder*. Further, the organizations discussed in the report each voiced concern over the lack of competition as a significant hindrance to the effective application of auctions.

In this paper, we develop a new class of mechanisms, *robust mechanisms*, that allow for some degree of uncertainty in the distribution over the bidders types while still performing better than weakly prior-dependent mechanisms such as ex-post mechanisms. We first start with a negative result: *that no robust mechanism can guarantee performance any better than an ex-post mechanism.* However, we also provide an algorithm for computing a class of robust mechanisms that we refer to as *robust automated mechanisms* in polynomial time in the number of bidder types. This mechanism design algorithm is primarily targeted at applications with few bidders, or thin markets, due to exponential scaling in the number of bidders. This mechanism design technique combines ideas from the automated mechanism design literature [Conitzer and Sandholm 2007; Conitzer et al. 2007; Guo and Conitzer 2010] and the literature on robust optimization [Bertsimas and Sim 2004; Aghassi and Bertsimas 2006]. We then introduce the notion of *ϵ-robust mechanisms*, or mechanisms that guarantee with high probability that the standard constraints of incentive compatibility and individual rationality hold but allow for a non-zero chance of violation. Finally, we show experimentally that ϵ-robust mechanisms can significantly outperform other mechanism design procedures when the mechanism designer can estimate the distribution over a bidder's type and a correlated external signal.

2. PRELIMINARIES

We consider a single monopolistic seller auctioning one object, which the seller values at zero, to a single bidder whose valuation is correlated with an external signal. The special case of a single bidder and an externally verifiable signal captures many of the important aspects of this problem and while increasing ease of exposition relative to the case of many bidders, and this setting has been used in the literature on correlated mechanism design [McAfee and Reny 1992; Albert et al. 2015, 2016] for this purpose. The external signal can, but does not necessarily, represent other bidders' bids. For the case where the mechanism is guaranteed to be incentive compatible and individually rational with probability one, this is without loss of generality as our results can easily be extended to include multiple bidders. If the mechanism fails to be incentive compatible or individually rational with positive probability (as will be the case in Section 4), this is not without loss of generality. However, it allows us to ignore the situation where the external signal is intentionally mis-reported, and this significantly simplifies the problem. We have defined strategies that we believe will overcome this limitation, but this remains a topic for future work.

The bidder has a type θ drawn from a finite set of discrete types $\Theta = \{1, ..., |\Theta|\}$. Further, the bidder has a valuation function $v : \Theta \rightarrow \mathbb{R}^+$ that maps types to valuations for the object. Assume without loss of generality that for all i and $\theta, \theta' \in \Theta$, if $\theta > \theta'$ then $v(\theta) > v(\theta')$. The discrete external signal is denoted by $\omega \in \Omega = \{1, 2, \ldots, |\Omega|\}$. Throughout the paper, we will denote vectors as bold symbols.

There is a probability distribution, $\boldsymbol{\pi}$, over the types of the bidder and external signal where the probability of type and signal (θ, ω) is $\boldsymbol{\pi}(\theta, \omega)$. The probability distribution can be represented in many possible ways, but we will represent it as a matrix. Specifically, the distribution is a matrix of dimension $|\Theta| \times |\Omega|$ whose elements are all positive and sum to one. Note that in contrast to much of the literature on mechanism design, we do not require that the bidder type be distributed independently of the external signal.

The distribution over the external signal ω given θ will be denoted by the $|\Omega|$ dimensional vector $\boldsymbol{\pi}(\bullet|\theta)$. We will, in many cases, be primarily interested in the conditional distribution over the external signal given the bidder's type, $\boldsymbol{\pi}(\bullet|\theta)$, so we will represent the full distribution as a marginal distribution over Θ, $\boldsymbol{\pi}_\theta$, and a set of conditional distributions over Ω, $\boldsymbol{\pi}(\bullet|\bullet) = \{\boldsymbol{\pi}(\bullet|1), \boldsymbol{\pi}(\bullet|2), ..., \boldsymbol{\pi}(\bullet||\Theta|)\}$. Therefore, if the true distribution is $\boldsymbol{\pi}$, we will alternatively represent it as $\{\boldsymbol{\pi}_\theta, \boldsymbol{\pi}(\bullet|\bullet)\}$.

A (direct) revelation mechanism is defined by, given the bidder type and external signal (θ, ω), 1) the probability that the seller allocates the item to the bidder and 2) a monetary transfer from the bidder to the seller. We will denote the probability of allocating the item to the bidder as $\boldsymbol{p}(\theta, \omega)$, which is a value between zero and one, and the transfer from the bidder to the seller as $\boldsymbol{x}(\theta, \omega)$, where a positive value denotes a payment to the seller and a negative value a payment from the seller to the bidder. We will denote the set of allocation probabilities and payments as $(\boldsymbol{p}, \boldsymbol{x})$.

Definition 2.1 (*Bidder's Utility*). Given a realization of the external signal ω, reported type $\theta' \in \Theta$ by the bidder, and true type $\theta \in \Theta$, the bidder's utility under mechanism $(\boldsymbol{p}, \boldsymbol{x})$ is:

$$U(\theta, \theta', \omega) = v(\theta)p(\theta', \omega) - x(\theta', \omega) \tag{1}$$

Due to the well-known revelation principle [Gibbons 1992], the seller can restrict her attention to incentive compatible mechanisms, i.e., mechanisms where it is always optimal for the bidder to truthfully report his valuation. However, incentive compatibility can be specified in multiple ways. For the sake of presentation, we will restrict our focus to two of the most common, *ex-post incentive compatibility* and *Bayesian incentive compatibility*. Ex-post incentive compatible mechanisms guarantee that for any realization of the external signal, the bidder always finds it optimal to report his value truthfully. In contrast, Bayesian incentive compatible mechanisms only guarantee that, given the beliefs of the bidder over the external signal, the bidder will have the highest *expected* utility if he reports truthfully: after seeing the realization of the external signal, he may regret his report.

Definition 2.2 (*Ex-Post Incentive Compatibility*). A mechanism $(\boldsymbol{p}, \boldsymbol{x})$ is *ex-post incentive compatible (IC)* if:

$$\forall \theta, \theta' \in \Theta, \omega \in \Omega : U(\theta, \theta, \omega) \geqslant U(\theta, \theta', \omega) \tag{2}$$

Definition 2.3 (*Bayesian Incentive Compatibility*). A mechanism $(\boldsymbol{p}, \boldsymbol{x})$ is *Bayesian incentive compatible (IC)* if:

$$\forall \theta, \theta' \in \Theta : \sum_{\omega \in \Omega} \boldsymbol{\pi}(\omega|\theta) U(\theta, \theta, \omega) \geqslant \sum_{\omega \in \Omega} \boldsymbol{\pi}(\omega|\theta) U(\theta, \theta', \omega) \tag{3}$$

Bayesian incentive compatibility is a statement about the *beliefs* of the bidder over the external signal, $\boldsymbol{\pi}(\omega|\theta)$. Specifically, it allows the seller to determine payments by *lottery*. The lottery that bidder i faces can be dependent on his valuation, but the lottery itself is over the external signal. In order for the mechanism to be incentive compatible the bidder must *believe* that his expected utility is higher from the lottery he gets by reporting his valuation truthfully than by reporting any other valuation (see Albert et al. 2016 for an in depth exploration of this point). Bayesian incentive compatibility is a strict relaxation of ex-post in the sense that any mechanism that is ex-post incentive compatible is also Bayesian incentive compatible.

In addition to incentive compatibility, we are interested in mechanisms that are *individually rational*, i.e. it is rational for a bidder to participate in the mechanism. We will define *ex-post individual rationality* (a bidder is *never* worse off by participating in the mechanism) and *ex-interim individual rationality* (the bidder has non-zero *expected* utility for participating in the mechanism). Again, ex-interim individual rationality is a strict relaxation of ex-post.

Definition 2.4 (*Ex-Post Individual Rationality*). A mechanism $(\boldsymbol{p}, \boldsymbol{x})$ is *ex-post individually rational (IR)* if:

$$\forall \theta \in \Theta, \omega \in \Omega : U(\theta, \theta, \omega) \geqslant 0 \tag{4}$$

Definition 2.5 (*Ex-Interim Individual Rationality*). A mechanism $(\boldsymbol{p}, \boldsymbol{x})$ is *ex-interim individually rational (IR)* if:

$$\forall \theta \in \Theta : \sum_{\omega \in \Omega} \boldsymbol{\pi}(\omega|\theta) U(\theta, \theta, \omega) \geqslant 0 \tag{5}$$

We will refer to mechanisms that satisfy ex-post individual rationality and incentive compatibility as *ex-post mechanisms* and mechanisms that satisfy Bayesian incentive compatibility and ex-interim individual rationality as *Bayesian mechanisms*. Bayesian mechanisms are what we have been referring to as prior-dependent mechanisms, while ex-post is weakly prior-dependent, i.e. only the objective function depends on the distribution, not the constraints over incentive compatibility and individual rationality.

To illustrate the importance of prior-dependent mechanisms, it is necessary to review a few important results in the literature on revenue maximization with correlated valuation distributions when the distribution is perfectly known.

Definition 2.6 (*Cremer-McLean Condition*). The distribution over types $\boldsymbol{\pi}(\bullet)$, is said to satisfy the Cremer-McLean condition if the set of beliefs associated with the bidder, $\{\boldsymbol{\pi}(\bullet|\theta) : \theta \in \Theta\}$, are linearly independent.

THEOREM 2.7 (CREMER AND MCLEAN 1985). *If the Cremer-McLean condition is satisfied by the distribution $\boldsymbol{\pi}(\bullet)$, then there exists an ex-interim IR and ex-post IC mechanism that extracts the full social surplus as revenue.*

This result due to Cremer and McLean 1985 states that under the apparently reasonable Cremer-McLean condition, i.e. a condition that holds with probability one for a random distribution, the mechanism designer can generate as much revenue in expectation as if she knew the bidder's valuation. This is a remarkable result and it can be relaxed further, for both ex-post and Bayesian IC, by the results in Albert et al. 2016.

THEOREM 2.8 (ALBERT ET AL. 2016). *A Bayesian IC and ex-interim IR mechanism can extract full social surplus as revenue if and only if there exists a* convex *function $G : \mathbb{R}^{|\Omega|} \to \mathbb{R}$ such that for all $\theta \in \Theta$, $G(\boldsymbol{\pi}(\bullet|\theta)) = -v(\theta)$.*

3. STRICTLY BOUNDED ESTIMATION OF THE DISTRIBUTION

While Theorems 2.7 and 2.8 make relatively weak assumptions about the distributions in order to guarantee full revenue extraction, they do require that the mechanism designer knows the distribution exactly. If instead of precise knowledge of the distribution of bidder types and external signals the mechanism designer has an imprecise estimate of the distribution, the prior-dependent mechanism can fail to be both incentive compatible and individually rational. This failure can be a significant problem for two reasons. First, if the mechanism is not individually rational bidders will not participate in the mechanism. If the market is thin (or as in our setting, a single bidder), the loss of even a single bidder can lead to significant decreases in expected revenue, even relative to simple mechanisms [Bulow and Klemperer 1996]. Second, if the mechanism is not incentive compatible, the bidder may optimally choose to misreport his true valuation, leading both to biases in future estimates of the distribution and difficulty in reasoning about the performance of the mechanism, since it is unclear a-priori how the bidder will report.

It is in this sense that Bayesian incentive compatible and ex-interim individually rational mechanisms are, in general, strongly prior-dependent. The mechanism depends not only on the seller's estimate of the distribution, but also the bidder's belief over the distribution. The consequences of these being mis-aligned is not just slightly lower expected revenue, as would be the case for weakly prior-dependent mechanisms such as a second price auction with reserve; it is a failure of the mechanism to maintain fundamental characteristics [Hartline 2014; Albert et al. 2015]. Therefore, unless the seller has perfect knowledge of the bidder's beliefs, standard mechanism design techniques will leave only the option of using sub-optimal, weakly prior-dependent mechanisms.

A more realistic assumption is that the distribution is not perfectly known, but instead estimated. The seller estimates the distribution $\boldsymbol{\pi}$ as $\hat{\boldsymbol{\pi}}$. Assume that this estimation is imperfect, and that there exists a set of distributions that are consistent with the estimated distribution.

Definition 3.1 (*Set of Consistent Distributions*). Let $P(A)$ be the set of probability distributions over A. Then the space of all probability distributions over $\Theta \times \Omega$ can be represented as the Cartesian product $P(\Theta) \times \prod_{\omega \in \Omega} P(\Omega)$. A subset $\mathcal{P}(\hat{\boldsymbol{\pi}}) = \mathcal{P}(\{\hat{\boldsymbol{\pi}}_\theta, \hat{\boldsymbol{\pi}}(\bullet|\bullet)\}) \subseteq P(\Theta) \times \prod_{\omega \in \Omega} P(\Omega)$ is a *consistent set of distributions* for the estimated distribution $\{\hat{\boldsymbol{\pi}}_\theta, \hat{\boldsymbol{\pi}}(\bullet|\bullet)\}$ if the true distribution, $\{\boldsymbol{\pi}_\theta, \boldsymbol{\pi}(\bullet|\bullet)\}$, is guaranteed to be in $\mathcal{P}(\hat{\boldsymbol{\pi}})$.

With a consistent set of distributions, we can relax the notion of ex-interim IR and Bayesian IC by requiring that the mechanism be IR and IC for all distributions in the consistent set. However, since the distribution $\boldsymbol{\pi}$ is also private information, by the revelation principle, the mechanism designer can also the true distribution and condition the mechanism on the reported distribution. Therefore, we modify the definitions of the mechanism, $(\boldsymbol{p}, \boldsymbol{x})$, such that they depend not only on the reported type and external signal, but also the reported distribution $\boldsymbol{\pi}'$. We similarly modify the definition of bidder utility.

Definition 3.2 (*Robust Individual Rationality*). A mechanism is *robust individually rational* for estimated bidder distribution $\hat{\boldsymbol{\pi}}$ and consistent set of distributions

$\mathcal{P}(\hat{\boldsymbol{\pi}})$ if for all $\theta \in \Theta$ and $\boldsymbol{\pi} \in \mathcal{P}(\hat{\boldsymbol{\pi}})$,

$$\sum_{\omega \in \Omega} \boldsymbol{\pi}(\omega|\theta) U(\theta, \boldsymbol{\pi}, \theta, \boldsymbol{\pi}, \omega) \geqslant 0 \tag{6}$$

Definition 3.3 (*Robust Incentive Compatibility*). A mechanism is *robust incentive compatible* for estimated bidder distribution $\hat{\boldsymbol{\pi}}$ and consistent set of distributions $\mathcal{P}(\hat{\boldsymbol{\pi}})$ if for all $\theta, \theta' \in \Theta$ and $\boldsymbol{\pi}, \boldsymbol{\pi}' \in \mathcal{P}(\hat{\boldsymbol{\pi}})$,

$$\sum_{\omega \in \Omega} \boldsymbol{\pi}(\omega|\theta) U(\theta, \boldsymbol{\pi}, \theta, \boldsymbol{\pi}, \omega) \geqslant \sum_{\omega \in \Omega} \boldsymbol{\pi}(\omega|\theta) U(\theta, \boldsymbol{\pi}, \theta', \boldsymbol{\pi}', \omega) \tag{7}$$

3.1. Inapproximability of Optimal Mechanism

It is natural to ask whether or not we can, with a sufficiently precise estimate of the true distribution, expect to do approximately as well as the case of perfect knowledge of the true distribution. Previous work [Fu et al. 2014] has shown that with a discrete set of distributions and the ability to sample from the underlying distribution, with a relatively small number of samples (specifically a number of samples on the order of the number of discrete distributions) the mechanism designer can extract full surplus as revenue. This previous work relied heavily on the requirement of a finite number of distributions, and in our setting we allow for an infinite number of distributions and continuous sets of consistent distributions. The following example demonstrates that in this context it is, unfortunately, impossible to guarantee a good approximation to the revenue achievable with perfect knowledge of the distribution. In demonstrating this, we will require Lemma 3.4.

LEMMA 3.4 (MARTIN 1975). *Let $\boldsymbol{a}(\boldsymbol{t}), \boldsymbol{b}(\boldsymbol{t}), \boldsymbol{c}(\boldsymbol{t})$, and $\boldsymbol{d}(\boldsymbol{t})$ be vectors parameterized by the parameter vector $\boldsymbol{t} \in \mathcal{Q}$. Assume that $\boldsymbol{a}(\boldsymbol{t}), \boldsymbol{b}(\boldsymbol{t}), \boldsymbol{c}(\boldsymbol{t})$, and $\boldsymbol{d}(\boldsymbol{t})$ converge continuously to $\boldsymbol{a}(\boldsymbol{0}), \boldsymbol{b}(\boldsymbol{0}), \boldsymbol{c}(\boldsymbol{0})$, and $\boldsymbol{d}(\boldsymbol{0})$ as $\boldsymbol{t} \to \boldsymbol{0}$. Similarly, $A(\boldsymbol{t}), B(\boldsymbol{t}), C(\boldsymbol{t})$, and $D(\boldsymbol{t})$ are matrices that converge continuously to $A(\boldsymbol{0}), B(\boldsymbol{0}), C(\boldsymbol{0})$, and $D(\boldsymbol{0})$.*

Define the parameterized linear program $LP(\boldsymbol{t})$ as:

$$\max_{\boldsymbol{x}, \boldsymbol{q}} \quad \boldsymbol{c}'(\boldsymbol{t})\boldsymbol{x} + \boldsymbol{d}'(\boldsymbol{t})\boldsymbol{q} \tag{8}$$

subject to

$$A(\boldsymbol{t})\boldsymbol{x} + B(\boldsymbol{t})\boldsymbol{q} = \boldsymbol{a}(\boldsymbol{t}) \tag{9}$$

$$C(\boldsymbol{t})\boldsymbol{x} + D(\boldsymbol{t})\boldsymbol{q} \leqslant \boldsymbol{b}(\boldsymbol{t}) \tag{10}$$

$$\boldsymbol{q} \geqslant \boldsymbol{0} \tag{11}$$

Then if the set of optimal solutions of $LP(\boldsymbol{0})$, $\{(\boldsymbol{x}, \boldsymbol{q}) : (\boldsymbol{x}, \boldsymbol{q}) \in \arg\max(LP(\boldsymbol{0}))\}$, is bounded, then the objective value of $LP(\boldsymbol{t})$ is upper semi-continuous at $\boldsymbol{t} = \boldsymbol{0}$.

Example 3.5. Suppose that the marginal distribution over the bidder type is exactly known, but the conditional distribution over the external signal given the bidders type is restricted to a set of consistent distributions. Let the marginal distribution over the type of the bidder be given by $\pi(\theta) = 1/2^\theta$ for $\theta \in \{1, ..., |\Theta|-1\}$ and $\pi(|\Theta|) = 1/2^{|\Theta|-1}$. Further let the value of the bidder for the item be $v(\theta) = 2^\theta$. Therefore, the expected value of the bidder's valuation is $\sum_{\theta=1}^{|\Theta|-1} (1/2^\theta) 2^\theta + (1/2)^{|\Theta|-1} 2^{|\Theta|} = |\Theta| + 1$.

We will assume that the external signal is binary, i.e. $\Omega = \{1, 2\}$. Suppose that the beliefs of the bidder are such that $\pi(1|\theta) = 1 - \pi(2|\theta) \in [1/2 - \epsilon, 1/2 + \epsilon]$ where $\epsilon \in (0, 1/2]$. Therefore, for $\theta \in \{1, ..., |\Theta|-1\}$, $\pi(\theta, \omega) = \pi(\omega|\theta)\pi(\theta) \in [1/2^{\theta+1} - (1/2^\theta)\epsilon, 1/2^{\theta+1} + (1/2^\theta)\epsilon]$, and for $\pi(|\Theta|, \omega) = \pi(\omega|\theta)\pi(|\Theta|) \in [1/2^{|\Theta|} - (1/2^{|\Theta|-1})\epsilon, 1/2^{|\Theta|} + (1/2^{|\Theta|-1})\epsilon]$. We will refer to the estimated distribution associated with this consistent set as $\hat{\boldsymbol{\pi}}_\epsilon$, and the consistent set of distributions such that for all $\theta \in \Theta$ and $\omega \in \Omega$, $\pi(1|\theta) = 1 - \pi(2|\theta) \in [1/2 - \epsilon, 1/2 + \epsilon]$ as the set $\mathcal{P}(\hat{\boldsymbol{\pi}}_\epsilon)$.

We will be particularly interested in two sets of special distributions consistent with the estimate distribution. Specifically, we will refer to the set of distributions such that for all $\theta \in \Theta$, $\pi(1|\theta) = \gamma$ where $\gamma \in [1/2 - \epsilon, 1/2 + \epsilon]$ as the set of independent private value (IPV) distributions, and we will refer to a particular element of the set as $IPV(\gamma)$. For all distributions in the IPV set, the bidder's valuation is uncorrelated with the external signal, and standard results from independent valuation settings apply [Myerson 1981]. We will additionally refer to the set of distributions such that $\pi(1|\theta) = \gamma + \kappa(1 - (1/2)^{|\Theta| - \theta})$ where $\gamma \in [1/2 - \epsilon, 1/2 + \epsilon]$ and $\kappa \in [1/2 - \epsilon - \gamma, 0) \cup (0, 1/2 + \epsilon - \gamma]$. We will refer to this set as the full extraction set (FE) and an element of this set as $FE(\gamma, \kappa)$, and by the definition of γ and κ, $IPV(\gamma), FE(\gamma, \kappa) \in \mathcal{P}(\hat{\boldsymbol{\pi}}_\epsilon)$.

Note that this example is a special case of the definition of consistent distributions given in Definition 4.1. Specifically, in this example, the beliefs of the bidder are uncertain, but the marginal distribution over the bidder's types is perfectly known. Therefore, when we refer to a consistent distribution given the estimate of the true distribution, we mean a distribution such that the marginal distribution is identical to that specified above and the belief of the bidder is consistent with the estimate. Given that this setting is strictly more restrictive than the general problem, we can generalize to the less restrictive setting.

LEMMA 3.6. *In the setting of Example 3.5, for all distributions $FE(\gamma, \kappa)$, there exists a Bayesian IC and ex-interim IR mechanism such that the expected revenue is $|\Theta| + 1$.*

PROOF. Let $\boldsymbol{\pi}(\bullet) = FE(\gamma, \kappa)$. Also, define the linear function $G(\boldsymbol{\pi}(\bullet|\theta)) = \boldsymbol{\pi}(1|\theta)2^{|\Theta|}/\kappa - \gamma 2^{|\Theta|}/\kappa + 2^{|\Theta|}$. Then, $G(\boldsymbol{\pi}(\bullet|\theta)) = 2^\theta$, and by Theorem 2.8, there exists a mechanism that extracts the full surplus as revenue, and therefore has an expected value equal to the expectation of the bidders valuation, $|\Theta| + 1$. □

LEMMA 3.7. *For all $\boldsymbol{\pi} \in \mathcal{P}(\hat{\boldsymbol{\pi}}_\epsilon)$, there exists a mechanism with expected revenue 2. Further, there exists a distribution $\boldsymbol{\pi}' \in \mathcal{P}(\hat{\boldsymbol{\pi}}_\epsilon)$ such that the optimal Bayesian IC and ex-interim IR mechanism has an expected revenue 2.*

PROOF. First, a take it or leave it offer of $2^{|\Theta|}$ has an expected value of 2 for all distributions $\boldsymbol{\pi} \in \mathcal{P}(\hat{\boldsymbol{\pi}}_\epsilon)$. Second, let $\pi'(\bullet) = IPV(\gamma)$. Then θ and ω are independent and the optimal mechanism is a take it or leave it offer [Myerson 1981]. It is trivial to verify by direct calculation that the maximum revenue achievable, 2, from a take it or leave it offer for this setting is when the price of the item is set to $2^{|\Theta|}$, and for any other price for the take it or leave it offer the expected revenue is strictly less than 2. □

LEMMA 3.8. *For the setting of Example 3.5, any mechanism that is robust IC and robust IR over the set of distributions $\mathcal{P}(\hat{\boldsymbol{\pi}}_\epsilon)$, and that guarantees revenue of at least 2 for all distributions, for all distributions $\boldsymbol{\pi} \in \mathcal{P}(\hat{\boldsymbol{\pi}}_{\epsilon/2})$, is bounded, i.e. for all $\theta \in \Theta$ and $\omega \in \Omega$, there exists an $M \in \mathbb{R}$ such that $|\boldsymbol{x}(\theta, \omega, \boldsymbol{\pi})| < M$.*

PROOF. Suppose not, then for all $M \in \mathbb{R}$, there exists some distribution $\boldsymbol{\pi} \in \mathcal{P}(\hat{\boldsymbol{\pi}}_{\epsilon/2})$ and $\theta \in \Theta$ and $\omega \in \Omega$ such that $|\boldsymbol{x}(\theta, \omega, \boldsymbol{\pi})| > M$. Assume without loss of generality that $\omega = 1$. There are two cases, either 1) $\boldsymbol{x}(\theta, 1, \boldsymbol{\pi}) > M$ or 2) $\boldsymbol{x}(\theta, 1, \boldsymbol{\pi}) < -M$. Consider case 2. Choose $M = (2^{|\Theta|+3}/\epsilon) * (|\Theta| + 1)$. Then $\boldsymbol{x}(\theta, 2, \boldsymbol{\pi}) > 0$ or the mechanism designer cannot earn positive revenue, a contradiction. Let $\boldsymbol{\pi}' = IPV(1/2 + \epsilon)$. Then, by robust

IC for $\theta' \in \Theta$ and IR for θ and $\boldsymbol{\pi}$, **Michael: Need to provide more dtails.**

$$\begin{aligned}
\sum_{\omega} \boldsymbol{\pi}'(\omega|\theta')U(\theta', \boldsymbol{\pi}', \theta', \boldsymbol{\pi}', \omega) \geqslant \sum_{\omega} \boldsymbol{\pi}'(\omega|\theta')U(\theta', \boldsymbol{\pi}', \theta, \boldsymbol{\pi}, \omega) - \sum_{\omega} \boldsymbol{\pi}(\omega|\theta)U(\theta, \boldsymbol{\pi}, \theta, \boldsymbol{\pi}, \omega)\\
&= (\boldsymbol{\pi}'(1|\theta') - \boldsymbol{\pi}(1|\theta)) * (-\boldsymbol{x}(\theta, 1, \boldsymbol{\pi})) + (v(\theta')\boldsymbol{\pi}'(1|\theta') - v(\theta)\boldsymbol{\pi}(1|\theta))\boldsymbol{p}(\theta, 1, \boldsymbol{\pi})\\
&\quad + (\boldsymbol{\pi}'(2|\theta') - \boldsymbol{\pi}(2|\theta))) * (-\boldsymbol{x}(\theta, 2, \boldsymbol{\pi})) + (v(\theta')\boldsymbol{\pi}'(2|\theta') - v(\theta)\boldsymbol{\pi}(2|\theta))\boldsymbol{p}(\theta, 2, \boldsymbol{\pi})\\
&\geqslant (\boldsymbol{\pi}'(1|\theta') - \boldsymbol{\pi}(1|\theta)) * (-\boldsymbol{x}(\theta, 1, \boldsymbol{\pi})) - 2^{|\Theta|+1}\\
&\geqslant (\epsilon/2) * (-\boldsymbol{x}(\theta, 1, \boldsymbol{\pi})) - 2^{|\Theta|+1} > 2^{|\Theta|}.
\end{aligned}$$

Therefore, the mechanism cannot extract positive revenue from a bidder with distribution $\boldsymbol{\pi}'$, a contradiction to the assumption that the mechanism guarantees revenue of at least 2 for all distributions in $\mathcal{P}(\hat{\boldsymbol{\pi}}_\epsilon)$.

Now suppose that $\boldsymbol{x}(\theta, 1, \boldsymbol{\pi}) > M$. Since we know that there is some M' such that for all $\theta \in \Theta$, $\omega \in \Omega$, and $\boldsymbol{\pi} \in \mathcal{P}(\hat{\boldsymbol{\pi}}_{\epsilon/2})$, $\boldsymbol{x}(\theta, 1, \boldsymbol{\pi}) \geqslant -M'$, choose $M = 2^{|\Theta|} + (1/2 + \epsilon/2)(2^{|\Theta|} + M')/(1/2 - \epsilon/2)$. Then,

$$\begin{aligned}
\sum_{\omega} \boldsymbol{\pi}(\omega|\theta)U(\theta, \boldsymbol{\pi}, \theta, \boldsymbol{\pi}, \omega) &= \boldsymbol{\pi}(1|\theta) * (v(\theta)\boldsymbol{p}(\theta, 1, \boldsymbol{\pi}) - \boldsymbol{x}(\theta, 1, \boldsymbol{\pi}))\\
&\quad + \boldsymbol{\pi}(2|\theta) * (v(\theta)\boldsymbol{p}(\theta, 2, \boldsymbol{\pi}) - \boldsymbol{x}(\theta, 2, \boldsymbol{\pi}))\\
&\leqslant (1/2 - \epsilon/2) * (2^{|\Theta|} - \boldsymbol{x}(\theta, 1, \boldsymbol{\pi})) + (1/2 + \epsilon/2) * (2^{|\Theta|} + M')\\
&< 0.
\end{aligned}$$

This is a contradiction of individual rationality. □

THEOREM 3.9. *For all $\delta > 0$ and all robust mechanisms $(\boldsymbol{p}, \boldsymbol{x})$ for the setting of example 3.5, there exists a distribution $\boldsymbol{\pi}' \in \mathcal{P}(\hat{\boldsymbol{\pi}}_\epsilon)$ such that the Bayesian IC and Ex-Interim IR mechanism achieves revenue of $|\Theta| + 1$ while the robust mechanism achieves revenue of 2.*

PROOF. Suppose not. Suppose that there exists a robust mechanism $(\boldsymbol{p}, \boldsymbol{x})$ such that for all distributions for which a Bayesian IC and Ex-Interim IR mechanism achieves full surplus extraction, the robust mechanism achieves revenue strictly greater than 2. Consider the distribution $\boldsymbol{\pi} = FE(1/2, \epsilon/2)$. Then there must exist $\theta^* \in \Theta$ where $\theta^* < |\Theta|$ such that the item is allocated to the bidder with positive probability, otherwise the maximum revenue that could be extracted for the distribution $\boldsymbol{\pi}$ is $\pi_\theta(|\Theta|)2^{|\Theta|} = 2$ by individual rationality. Now consider the distribution $\boldsymbol{\pi}' = IPV(\boldsymbol{\pi}(1|\theta^*))$. By robust incentive compatibility,

$$\sum_{\omega} \boldsymbol{\pi}'(\omega|\theta)U(|\Theta|, \boldsymbol{\pi}', |\Theta|, \boldsymbol{\pi}', \omega) \geqslant \sum_{\omega} \boldsymbol{\pi}'(\omega|\theta)U(|\Theta|, \boldsymbol{\pi}', \theta^*, \boldsymbol{\pi}, \omega) > 0$$

It is easy to verify by direct application of Bayesian incentive compatibility and ex-interim individual rationality, that the maximum revenue achievable for $IPV(\boldsymbol{\pi}(1|\theta^*))$ when $\sum_\omega \boldsymbol{\pi}'(\omega|\theta)U(|\Theta|, \boldsymbol{\pi}', |\Theta|, \boldsymbol{\pi}', \omega) > 0$, is strictly less than 2. Further, by lemma 3.8, the payments for any robust mechanism are strictly bounded by some $M \in \mathbb{R}$ for all distributions in $\mathcal{P}(\hat{\boldsymbol{\pi}}_{\epsilon/2})$. Therefore, the revenue for the distribution $IPV(\boldsymbol{\pi}(1|\theta^*))$ must be less than or equal to the objective value of the following linear program.

$$
\begin{aligned}
\max_{\boldsymbol{p},\boldsymbol{x}} & \sum_{\theta}\sum_{\omega}\boldsymbol{\pi}'(\theta,\omega)\boldsymbol{x}(\theta,\omega) \\
& \text{subject to} \\
& \sum_{\omega}\boldsymbol{\pi}'(\omega|\theta)U(\theta,\boldsymbol{\pi}',\theta,\boldsymbol{\pi}',\omega) \geqslant 0 \qquad \forall \quad \theta\in\Theta \\
& \sum_{\omega}\boldsymbol{\pi}'(\omega|\theta)U(\theta,\boldsymbol{\pi}',\theta,\boldsymbol{\pi}',\omega) \geqslant \sum_{\omega}\boldsymbol{\pi}'(\omega|\theta)U(\theta,\boldsymbol{\pi}',\theta,\boldsymbol{\pi}',\omega) \qquad \forall \quad \theta,\theta'\in\Theta \\
& \sum_{\omega}\boldsymbol{\pi}'(\omega|\theta)U(\theta,\boldsymbol{\pi}',\theta,\boldsymbol{\pi}',\omega) \geqslant \sum_{\omega}\boldsymbol{\pi}'(\omega|\theta)U(\theta,\boldsymbol{\pi}',\theta^*,\boldsymbol{\pi},\omega) \qquad \forall \quad \theta\in\Theta \\
& 0\leqslant \boldsymbol{p}(\theta,\omega)\leqslant 1 \qquad \forall \quad \theta\in\Theta,\omega\in\Omega \\
& -M\leqslant \boldsymbol{x}(\theta,\omega)\leqslant M \qquad \forall \quad \theta\in\Theta,\omega\in\Omega
\end{aligned}
\tag{12}
$$

Now, consider the series of distributions $\boldsymbol{\pi}_n = FE(\boldsymbol{\pi}(1|\theta^*), 1/n)$ where $n \in \mathbb{R}^+$. Assume that n is sufficiently large such that $\boldsymbol{\pi}_n \in \mathcal{P}(\hat{\boldsymbol{\pi}}_{\epsilon/2})$. Then the maximum revenue achievable by any robust mechanism on these distributions must also be less than or equal to the linear program defined by (12) with $\boldsymbol{\pi}'$ replaced by $\boldsymbol{\pi}_n$. Note that by Lemma 3.4, for any $\delta > 0$ there exists a $n \in \mathbb{R}^+$ such that the maximum revenue achievable for the distribution $\boldsymbol{\pi}_n$ must be less than or equal to δ more than the maximum revenue achievable for the distribution $IPV(\boldsymbol{\pi}(1|\theta))$, which is strictly less than 2. Therefore, there exists an $n \in \mathbb{R}$ such that the maximum revenue achievable is strictly less than 2, a contradiction. □

COROLLARY 3.10. *The expected revenue generated by a robust IR and robust IC mechanism guarantees at best an* $(|\Theta| + 1)/2$ *approximation to the revenue achievable by the optimal Bayesian IC and ex-interim IR mechanism if the distribution over types is exactly known.*

3.2. Optimal Robust Mechanism

Therefore, the worst case performance of any robust mechanism will be no better than that of an ex-post mechanism. However, there is potentially room to improve upon ex-post mechanisms with robust mechanisms, either by prioritizing certain segments of the set of consistent distributions that we are more likely to see, or by taking advantage of consistent distributions that do not contain an independent private value distribution.

However, existing mechanism design techniques are inadequate to optimize over these situations. Specifically, Bayesian mechanisms can have very unintuitive formulations consisting of multiple lotteries over the value of the external signal (see Albert et al. 2016 for a full discussion) making it unlikely that a standard class of simple intuitive mechanisms will be able to implement robust mechanisms. Therefore, we will combine techniques from automated mechanism design [Conitzer and Sandholm 2002, 2004; Guo and Conitzer 2010; Sandholm and Likhodedov 2015] and robust convex optimization [Bertsimas and Sim 2004; Aghassi and Bertsimas 2006] to automate the design of robust mechanisms.

Further, while it is theoretically possible to allow bidders to report both their valuations and their beliefs, and design optimal mechanisms given this joint report, standard automated mechanism design techniques require finite specified input, and we are explicitly interested in infinite sets of distributions. We will simplify the mechanism design process by only considering mechanisms for which the payments and probabilities of allocations depend on the reported bidder types and the realization of the external signal. While this is not without loss of generality, it will be sufficient to significantly outperform ex-post mechanisms in our experiments.

Definition 3.11 (*Optimal Robust Mechanism*). The *optimal robust mechanism* given an estimated distribution $\hat{\pi}$ and a consistent set of distributions $\mathcal{P}(\hat{\pi})$ is a mechanism that is an optimal solution to the following program:

$$\begin{aligned}
&\max_{\boldsymbol{x}(\theta,\omega),\boldsymbol{p}(\theta,\omega)} \sum_{\theta,\omega} \hat{\boldsymbol{\pi}}(\theta,\omega)\boldsymbol{x}(\theta,\omega) \\
&\text{subject to} \\
&\sum_{\omega\in\Omega} \boldsymbol{\pi}'(\omega|\theta)U(\theta,\theta,\omega) \geqslant 0 \qquad \forall \quad \theta\in\Theta, \boldsymbol{\pi}'(\bullet|\bullet)\in\mathcal{P}(\hat{\boldsymbol{\pi}}) \\
&\sum_{\omega\in\Omega} \boldsymbol{\pi}'(\omega|\theta)U(\theta,\theta,\omega) \geqslant \sum_{\omega\in\Omega} \boldsymbol{\pi}'(\omega|\theta)U(\theta,\theta',\omega) \qquad \forall \quad \theta,\theta'\in\Theta, \boldsymbol{\pi}'(\bullet|\bullet)\in\mathcal{P}(\hat{\boldsymbol{\pi}}) \\
&0 \leqslant \boldsymbol{p}(\theta,\omega) \leqslant 1 \qquad \forall \quad \theta\in\Theta, \omega\in\Omega
\end{aligned} \tag{13}$$

Note that the utility in (13) is a function of $(\boldsymbol{p},\boldsymbol{x})$. Notice also that the linear program defined by (13) still consists of an infinite number of constraints over a, potentially, non-convex set, and therefore is in general computationally infeasible. However, the following assumption allows computational tractability.

ASSUMPTION 3.12. *The set $\mathcal{P}(\hat{\boldsymbol{\pi}})$ can be characterized by the following: $\boldsymbol{\pi}'\in\mathcal{P}(\hat{\boldsymbol{\pi}})$ if and only if for all $\theta\in\Theta$ and $\omega\in\Omega$, $\pi'(\omega|\theta)\in[\underline{\boldsymbol{\pi}}(\omega|\theta),\overline{\boldsymbol{\pi}}(\omega|\theta)]$ for $0\leqslant\underline{\boldsymbol{\pi}}(\omega|\theta)\leqslant\overline{\boldsymbol{\pi}}(\omega|\theta)\leqslant 1$.*

We can relax Assumption 3.12 to allow for $\mathcal{P}(\hat{\pi})$ to be any set that can be characterized as being the interior of an $n-$polyhedron, where n is polynomial in the number of bidder types and external signals. However, for the sake of this work, we restrict our attention to the sets characterized by Assumption 3.12.

THEOREM 3.13. *For a given $(\boldsymbol{p},\boldsymbol{x})$ and $\mathcal{P}(\hat{\boldsymbol{\pi}})$ that satisfies Assumption 3.12, there exists a polynomial time algorithm that determines whether there exists a $\boldsymbol{\pi}'(\bullet|\theta)\in\mathcal{P}(\hat{\boldsymbol{\pi}}(\bullet|\theta))$ such that robust individual rationality is violated.*

PROOF. For each $\theta\in\Theta$, solve the following linear program

$$\begin{aligned}
&\min_{\boldsymbol{\pi}'(\omega|\theta)} \sum_{\omega} \boldsymbol{\pi}'(\omega|\theta)(v(\theta)\boldsymbol{p}(\omega,\theta)-\boldsymbol{x}(\omega,\theta)) \\
&\text{subject to} \\
&\boldsymbol{\pi}'(\omega|\theta) \geqslant \underline{\boldsymbol{\pi}}(\omega|\theta) \qquad \forall \quad \omega\in\Omega \\
&\boldsymbol{\pi}'(\omega|\theta) \leqslant \overline{\boldsymbol{\pi}}(\omega|\theta) \qquad \forall \quad \omega\in\Omega \\
&\sum_{\omega} \boldsymbol{\pi}'(\omega|\theta) = 1
\end{aligned} \tag{14}$$

Note that in the program (14), $(\boldsymbol{p},\boldsymbol{x})$ are no longer variables but coefficients. If (14) has an objective value of less than 0, then the robust IR constraint with distribution $\boldsymbol{\pi}'$ is violated. If the objective value is greater than 0, there is no robust IR constraint violated for θ.

There are $|\Theta|$ linear programs that must be solved. Each has a polynomial number of variables and constraints, so each linear program can be solved in polynomial time. Therefore, violated robust IR constraints can be generated in polynomial time. □

THEOREM 3.14. *For a given $(\boldsymbol{p}, \boldsymbol{x})$ and $\mathcal{P}(\hat{\boldsymbol{\pi}})$ that satisfies Assumption 3.12, there exists a polynomial time algorithm that determines whether there exists a $\boldsymbol{\pi}'(\bullet|\theta) \in \mathcal{P}(\hat{\boldsymbol{\pi}}(\bullet|\theta))$ such that robust incentive compatibility is violated.*

PROOF. For each $\theta, \theta' \in \Theta$, solve the following linear program

$$\min_{\boldsymbol{\pi}'(\omega|\theta)} \sum_{\omega} \boldsymbol{\pi}'(\omega|\theta)(v(\theta)\boldsymbol{p}(\omega,\theta) - \boldsymbol{x}(\omega,\theta) - (v(\theta)\boldsymbol{p}(\omega,\theta') - \boldsymbol{x}(\omega,\theta'))) \tag{15}$$

subject to

$$\boldsymbol{\pi}'(\omega|\theta) \geqslant \underline{\boldsymbol{\pi}}(\omega|\theta) \qquad \forall \quad \omega \in \Omega \tag{16}$$

$$\boldsymbol{\pi}'(\omega|\theta) \leqslant \overline{\boldsymbol{\pi}}(\omega|\theta) \qquad \forall \quad \omega \in \Omega \tag{17}$$

$$\sum_{\omega} \boldsymbol{\pi}'(\omega|\theta) = 1 \tag{18}$$

If the above program has an objective value of less than 0, then the IC constraint with distribution $\boldsymbol{\pi}'$ is violated. If the objective value is greater than 0, there is no robust IC constraint violated for (θ, θ').

There are $|\Theta|^2$ linear programs that must be solved. Each has a polynomial number of variables and constraints, so each linear program can be solved in polynomial time. Therefore, violated robust IC constraints can be generated in polynomial time. □

COROLLARY 3.15. *If $\mathcal{P}(\hat{\boldsymbol{\pi}})$ satisfies Assumption 3.12, the optimal robust mechanism can be calculated in a time polynomial in the number of types of the bidder and external signal.*

PROOF. By Theorems 3.13 and 3.14, we can determine whether or not a robust IR and robust IC constraint is violated in polynomial time, and add the constraint to the linear program. There are a $2|\Theta||\Omega|$ variables in the linear program, and there are $2|\Theta||\Omega|$ of non-IC and IR constraints.

Therefore, by the ellipsoid method, the optimal robust mechanism can be computed in polynomial time. □

Since for ex-post mechanisms, incentive compatibility and individual rationality are independent of the distribution, it would be expected that when we have no useful information about the distribution that robust mechanism design returns the optimal ex-post mechanism. The following corollary shows that this is indeed the case.

COROLLARY 3.16. *If for all $\theta \in \Theta$, and $\omega \in \Omega$, $\mathcal{P}(\hat{\boldsymbol{\pi}})$ is such that $\underline{\boldsymbol{\pi}}(\omega|\theta) = 0$ and $\overline{\boldsymbol{\pi}}(\omega|\theta) = 1$, then the optimal robust mechanism is equivalent to the optimal ex-post mechanism for the distribution $\hat{\boldsymbol{\pi}}$.*

PROOF. If for all $\theta \in \Theta$ and $\omega \in \Omega$, $\underline{\boldsymbol{\pi}}(\omega|\theta) = 0$ and $\overline{\boldsymbol{\pi}}(\omega|\theta) = 1$, then for each $\omega' \in \Omega$, the distribution such that $\boldsymbol{\pi}'(\omega|\theta) = 1$ if $\omega = \omega'$ and 0 otherwise is in $\mathcal{P}(\hat{\boldsymbol{\pi}})$. Therefore, the robust IR constraints contain the following set of constraints

$$v(\theta)p(\theta,\omega) - \boldsymbol{x}(\theta,\omega) \geqslant 0 \qquad \forall \quad \omega \in \Omega, \theta \in \Theta \tag{19}$$

which is ex-post IR.

Similarly, the robust IC constraints contain the following constraints

$$v(\theta)p(\theta,\omega) - \boldsymbol{x}(\theta,\omega) \geqslant v(\theta)p(\theta,\omega) - \boldsymbol{x}(\theta,\omega) \qquad \forall \quad \omega \in \Omega, \theta, \theta' \in \Theta \tag{20}$$

which is ex-post IC. □

4. ϵ-BOUNDED ESTIMATION OF THE DISTRIBUTION

While, so far we have been assuming that there is a well defined set, $\mathcal{P}(\hat{\boldsymbol{\pi}})$, such that the mechanism designer can guarantee that the true distribution, $\boldsymbol{\pi}$, is in the set, this is unlikely to be a realistic assumption in practice. It is far more reasonable that the mechanism designer would have a set such that the true distribution is in the set with some probability. If this is the case, we can still design mechanisms that are likely to outperform weakly prior dependent mechanisms, such as ex-post mechanisms, be relaxing the requirement that the mechanism be always incentive compatible and always individually rational. We will define the set of ϵ-consistent distributions as the following.

Definition 4.1 (*Set of ϵ-Consistent Distributions*). A subset $\mathcal{P}_\epsilon(\hat{\boldsymbol{\pi}}) = \mathcal{P}_\epsilon(\{\hat{\boldsymbol{\pi}}_\theta, \hat{\boldsymbol{\pi}}(\bullet|\bullet)\}) \subseteq P(\Theta) \times \prod_{\omega\in\Omega} P(\Omega)$ is an *ϵ-consistent set of distributions* for the estimated distribution $\{\hat{\boldsymbol{\pi}}_\theta, \hat{\boldsymbol{\pi}}(\bullet|\bullet)\}$ if the true distribution, $\{\boldsymbol{\pi}_\theta, \boldsymbol{\pi}(\bullet|\bullet)\}$, is in $\mathcal{P}_\epsilon(\hat{\boldsymbol{\pi}})$ with probability $1-\epsilon$.

Now we can define the notion of *ϵ-robust individual rationality* and *incentive compatibility*. These definitions are analogous to Definitions 3.2 and 3.3.

Definition 4.2 (*ϵ-Robust Individual Rationality*). A mechanism is *ϵ-robust individually rational* for estimated bidder distribution $\hat{\boldsymbol{\pi}}$ and ϵ-consistent set of distributions $\mathcal{P}_\epsilon(\hat{\boldsymbol{\pi}})$ if for all $\theta \in \Theta$ and $\boldsymbol{\pi} \in \mathcal{P}_\epsilon(\hat{\boldsymbol{\pi}})$,

$$\sum_{\omega\in\Omega} \boldsymbol{\pi}(\omega|\theta) U(\theta, \boldsymbol{\pi}, \theta, \boldsymbol{\pi}, \omega) \geqslant 0 \tag{21}$$

Definition 4.3 (*ϵ-Robust Incentive Compatibility*). A mechanism is *ϵ-robust incentive compatible* for estimated bidder distribution $\hat{\boldsymbol{\pi}}$ and ϵ-consistent set of distributions $\mathcal{P}_\epsilon(\hat{\boldsymbol{\pi}})$ if for all $\theta, \theta' \in \Theta$ and $\boldsymbol{\pi}, \boldsymbol{\pi}' \in \mathcal{P}_\epsilon(\hat{\boldsymbol{\pi}})$,

$$\sum_{\omega\in\Omega} \boldsymbol{\pi}(\omega|\theta) U(\theta, \boldsymbol{\pi}, \theta, \boldsymbol{\pi}, \omega) \geqslant \sum_{\omega\in\Omega} \boldsymbol{\pi}(\omega|\theta) U(\theta, \boldsymbol{\pi}, \theta', \boldsymbol{\pi}', \omega) \tag{22}$$

Again, for tractability of mechanism design, we will restrict attention to mechanisms that only depend on the reported bidder type and the external signal. This is not without loss of generality, but it greatly simplifies computation and implementation of the mechanism. As we will show in Section 5, we can still significantly outperform standard mechanism design techniques even with this restriction.

While we have focused on the restricted case of a single bidder and an external signal, all results up to this point readily generalize to multiple bidders with correlated valuations. However, in the case where a mechanism fails with positive probability to be incentive compatible, strategic interactions between bidders must be considered. Specifically, a bidder must consider whether or not another bidder is likely to misreport and affect the former bidder's optimal action. Restricting our focus to a single bidder allows us to assume that the bidder will always choose the utility maximizing action given his type, even if that action is not reporting his true type. Extending this to multiple bidders is a topic for future research.

Similarly to Definition 3.11, we can define the optimal optimal ϵ-robust mechanism.

Definition 4.4 (*Optimal ϵ-Robust Mechanism*). The *optimal ϵ-robust mechanism* given an estimated distribution $\hat{\boldsymbol{\pi}}$ and a ϵ-consistent set of distributions $\mathcal{P}_\epsilon(\hat{\boldsymbol{\pi}})$ is a

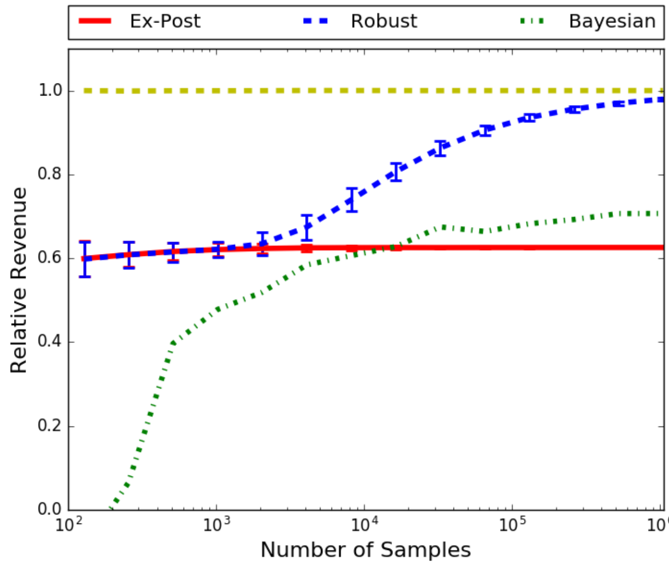


Fig. 1: The performance of the robust, ex-post, and Bayesian mechanisms using the estimated distribution. All revenue is scaled by the full social surplus, which is denoted as 1. Note that the Number of Samples is in log scale. The parameters used were as follows: Correlation = .5, ϵ = .05. Each experiment was repeated 200 times, and the 95% confidence interval is included for the robust and ex-post mechanisms. We do not include the confidence interval for the Bayesian mechanism due to it being too large to fit on the plot.

mechanism that maximizes the following program:

$$\max_{x(\theta,\omega),p(\theta,\omega)} \sum_{\theta,\omega} \hat{\boldsymbol{\pi}}(\theta,\omega)x(\theta,\omega)$$

subject to

$$\begin{aligned} &\sum_{\omega\in\Omega} \boldsymbol{\pi}'(\omega|\theta)U(\theta,\theta,\omega) \geqslant 0 \qquad \forall \quad \theta \in \Theta, \boldsymbol{\pi}'(\bullet|\bullet) \in \mathcal{P}_\epsilon(\hat{\boldsymbol{\pi}}(\bullet|\bullet)) \\ &\sum_{\omega\in\Omega} \boldsymbol{\pi}'(\omega|\theta)U(\theta,\theta,\omega) \geqslant \sum_{\omega\in\Omega} \boldsymbol{\pi}'(\omega|\theta)U(\theta,\theta',\omega) \qquad \forall \quad \theta,\theta' \in \Theta, \boldsymbol{\pi}'(\bullet|\bullet) \in \mathcal{P}_\epsilon(\hat{\boldsymbol{\pi}}(\bullet|\bullet)) \\ &0 \leqslant p(\theta,\omega) \leqslant 1 \qquad \forall \quad \theta \in \Theta, \omega \in \Omega \end{aligned} \tag{23}$$

COROLLARY 4.5. *If $\mathcal{P}_\epsilon(\hat{\boldsymbol{\pi}})$ satisfies Assumption 3.12, the optimal ϵ-robust mechanism can be calculated in a time polynomial in the number of bidder types and external signals.*

PROOF. The proof is identical to that of Corollary 3.15. □

5. EXPERIMENTAL RESULTS

While we've demonstrated that there can be no guarantee of full revenue extraction for all distributions, it is likely that under certain conditions, such as high correlation or precise estimates of the distribution, we can generate significantly more revenue using optimal ϵ-robust mechanisms. While it is ongoing work to characterize sufficient conditions for optimal ϵ-robust mechanisms to guarantee higher revenue than ex-post mechanisms under distribution uncertainty, we can gain some insight into the settings under which optimal ϵ-robust mechanism design techniques are likely to be the most successful.

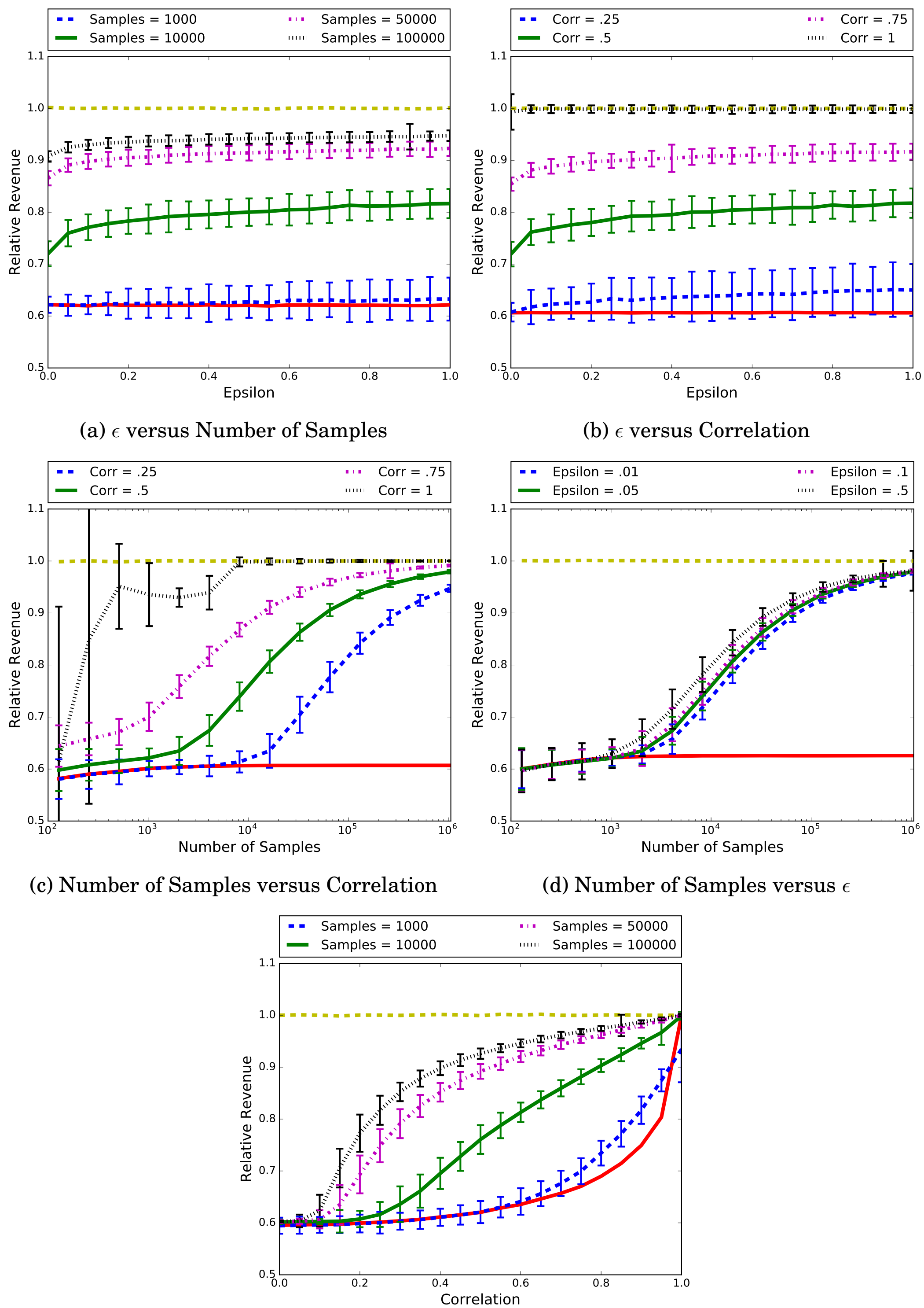
Throughout the experiments, we have a single bidder with $\theta \in \{1, 2, ..., 10\}$ types and valuation $v(\theta) = \theta$. The external signal is $\omega \in \{1, 2, ..., 10\}$. We model the true distribution as a categorical distribution with 10×10 elements, with each element corresponding to a tuple (θ, ω).

There are not, to our knowledge, standard distributions to test correlated mechanism design procedures available, so we used a discretized bi-variate normal distribution. Specifically, we discretized the area under the bi-variate standard normal distribution with varying correlation between $[-1.96, 1.96]$ in both dimensions as a 10×10 grid. We then assigned as the true probability for the tuple (i, j) the area under the probability density function of the bi-variate normal distribution in the square that is $i-1$ up from the bottom left corner (defined by $(-1.96, -1.96)$) and $j-1$ to the right. We then normalized the probability distribution over the tuples of bidder type and external signal. We chose the bi-variate normal distribution both for its broad relevance to many empirically observed distributions, but also for the ability to easily vary the correlation between the bidder type and the external signal. Note, also, that the bi-variate normal distribution always satisfies the Cremer-McLean condition if the correlation is positive.

To estimate the distribution, we sample from the true distribution and use Bayesian updating with a maximally uninformative Dirichlet prior ($\alpha = [1, ..., 1]$) to arrive at a Dirichlet posterior over the distribution of bidder types and external signals. We then calculate empirical confidence intervals by sampling from the Dirichlet posterior and observing the $\epsilon/(2 * 10 * 10)$ and $(1 - \epsilon/(2 * 10 * 10))$ quantiles for each element of the conditional distributions $\boldsymbol{\pi}(\omega|\theta)$ and used the quantiles as $\overline{\boldsymbol{\pi}}(\omega|\theta)$ and $\underline{\boldsymbol{\pi}}(\omega|\theta)$. Note that we do not use the $\epsilon/2$ and $(1 - \epsilon/2)$ quantiles due to jointly estimating confidence intervals for 100 variables. While the variables are not drawn independently of each other, we have done extensive testing and found that when we set the quantiles to $\epsilon/(2 * 10 * 10)$ and $(1 - \epsilon/(2 * 10 * 10))$, we observed almost exactly a proportion ϵ of the full distributions to have at least one element outside of the range $[\underline{\boldsymbol{\pi}}(\omega|\theta), \overline{\boldsymbol{\pi}}(\omega|\theta)]$ validating our choice of quantiles.

For our experiments, we solve for the optimal ex-post, ϵ-robust, and Bayesian mechanisms given our estimated distribution $\hat{\boldsymbol{\pi}}$ and our ϵ-consistent set of distributions $\mathcal{P}_\epsilon(\hat{\boldsymbol{\pi}})$. Given that both the optimal ϵ-robust and Bayesian mechanisms can fail to be both incentive compatible and individually rational due to the estimated distribution, we calculated the optimal action for the bidder to take when facing the mechanism: either report truthfully, strategically misreport, or do not participate in the mechanism. We then calculate the value of the mechanism as the probability of each bidder type in the true distribution times the revenue accrued by the seller when the bidder takes his optimal action under the true distribution.

In Figure 1, we show the performance of solving for the optimal ex-post, robust, and Bayesian mechanisms using our estimated distribution as we increase the number of samples. We report confidence intervals for both the ex-post mechanisms and the robust mechanisms; however for the Bayesian mechanisms, the confidence intervals were literally off the chart. Figure 1 demonstrates how badly the Bayesian mechanism performs when the distribution isn't exactly known. Even after $10,000$ samples from the true distribution, the Bayesian mechanism fails to outperform the ex-post mechanism. By contrast, the optimal ϵ-robust mechanism generates revenue indistinguishable from the ex-post mechanism for low numbers of samples, while significantly outperforming the ex-post mechanism starting at about $10,000$ samples.



(a) ϵ versus Number of Samples

(b) ϵ versus Correlation

(c) Number of Samples versus Correlation

(d) Number of Samples versus ϵ

(e) Correlation versus Number of Samples

Fig. 2: The performance of the robust and ex-post (red solid line) mechanisms using the estimated distribution. All revenue is scaled by the full social surplus, which is denoted as 1. Shown are the 95% confidence intervals for the robust mechanism repeated 200 times. For any variable not explicitly shown the following values were used: Number of Samples = 10000, Correlation = .5, ϵ = .05

In Figures 2a-2e, we show the effect of varying correlation, ϵ, and number of samples. We see that as the bidder type and external signal are more highly correlated, the ϵ-robust mechanism requires fewer samples to perform well, Figure 2c. Also, we see that the ϵ-robust mechanism is not very sensitive to the choice of ϵ, Figures 2a and 2d, a fact that we attribute to effectively being overly cautious in requiring all elements of the probability distribution to be in the bounded intervals. It is likely that even when a distribution is not exactly within the consistent set, the mechanism is still incentive compatible and individually rational. Further, with sufficient samples, even with an $\epsilon = 0$ we had fairly narrow confidence intervals due to the precision of the Dirichlet posterior.

Note that we consider the results here to be lower bounds on the performance of optimal ϵ-robust mechanisms. We assume a completely uninformative prior, increasing the required sample size. Further, we have used a naive distribution estimation procedure, so there is likely significant room to improve upon the estimation. All of these areas are ongoing work.

6. CONCLUSION AND FUTURE WORK

We have presented a new paradigm in mechanism design that formally addresses the problem of uncertainty in the bidder distribution. This is early work, and there is a large number of exciting directions to pursue. Specifically, while optimal robust mechanisms are trivially generalizable to multiple bidders, optimal ϵ-robust mechanisms are not. We are currently actively seeking to extend the mechanism design techniques presented here to assure that the mechanism remains incentive compatible and individually rational with probability $1 - \epsilon$ even in the case where there are multiple bidders, each choosing to mis-report or not participate as the bidder finds optimal.

In this work, we show that robust mechanism design (and therefore ϵ-robust mechanism design) cannot guarantee better performance than an ex-post mechanism for all possible consistent set of distributions. However, it seems likely that there are conditions under which a robust mechanism will be guaranteed to outperform an ex-post mechanism, specifically when the consistent set of distributions consists solely of distributions with large correlation between the bidder and the external signal. It would be very useful to characterize sufficient conditions on the consistent set of distributions and the valuation function of the bidder to guarantee higher revenue than all ex-post mechanisms.

While we demonstrate that with a continuous set of distributions there does not exist a robust mechanism that guarantees better performance, it is an open question as to whether there is a mechanism that incorporates sampling over the distribution that will guarantee increased revenue. This is known to be the case for a finite set of distributions [Fu et al. 2014], but for an infinite set of distributions, this is an open question.

Finally, there is much to be done experimentally to evaluate the performance of this ϵ-robust mechanism design procedure. Specifically, it may be optimal, when the number of samples from the true distribution is limited, to artificially restrict the number of external signals. Given that full revenue can be extracted under certain conditions even for a binary external signal [Albert et al. 2016], we conjecture that it may be possible to more efficiently use the samples by considering abstractions of the true distribution over the external signal.

REFERENCES

Michele Aghassi and Dimitris Bertsimas. 2006. Robust game theory. *Mathematical Programming* 107, 1 (2006), 231–273.

DOI:http://dx.doi.org/10.1007/s10107-005-0686-0

Michael Albert, Vincent Conitzer, and Giuseppe Lopomo. 2015. Assessing the Robustness of Cremer-McLean with Automated Mechanism Design. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI '15)*. AAAI Press, Austin, TX, USA, 763–769. http://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9296/9367

Michael Albert, Vincent Conitzer, and Giuseppe Lopomo. 2016. Maximizing Revenue with Limited Correlation: The Cost of Ex-Post Incentive Compatibility. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)*.

Dimitris Bertsimas and Melvyn Sim. 2004. The Price of Robustness. *Operations Research* 52, 1 (2004), 35–53. DOI:http://dx.doi.org/10.1287/opre.1030.0065

Jeremy Bulow and Paul Klemperer. 1996. Auctions Versus Negotiations. *The American Economic Review* 86, 1 (1996), pp. 180–194. http://www.jstor.org/stable/2118262

Vincent Conitzer and Tuomas Sandholm. 2002. Complexity of Mechanism Design. In *Proceedings of the 18th Annual Conference on Uncertainty in Artificial Intelligence (UAI '02)*. Morgan Kaufmann, Alberta, Canada, 103–110. http://dl.acm.org/citation.cfm?id=2073889

Vincent Conitzer and Tuomas Sandholm. 2004. Self-interested Automated Mechanism Design and Implications for Optimal Combinatorial Auctions. In *Proceedings of the 5th ACM Conference on Electronic Commerce (EC '04)*. ACM, New York, NY, USA, 132–141. DOI:http://dx.doi.org/10.1145/988772.988793

Vincent Conitzer and Tuomas Sandholm. 2007. Incremental Mechanism Design. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*. http://www.aaai.org/Papers/IJCAI/2007/IJCAI07-202.pdf

Vincent Conitzer, Tuomas Sandholm, and Craig Boutiller. 2007. Automated Design of Multistage Mechanisms. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*. http://www.cs.cmu.edu/~sandholm/multistageAMD.ijcai07.pdf

Jacques Cremer and Richard P. McLean. 1985. Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist when Demands are Interdependent. *Econometrica* 53, 2 (1985), pp. 345–361. http://www.jstor.org/stable/1911240

Jacques Cremer and Richard P. McLean. 1988. Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions. *Econometrica* 56, 6 (1988), pp. 1247–1257. http://www.jstor.org/stable/1913096

Hu Fu, Nima Haghpanah, Jason Hartline, and Robert Kleinberg. 2014. Optimal Auctions for Correlated Buyers with Sampling. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation (EC '14)*. ACM, New York, NY, USA, 23–36. DOI:http://dx.doi.org/10.1145/2600057.2602895

Robert Gibbons. 1992. *Game Theory for Applied Economists*. Princeton University Press.

Government Accountability Office. 2013. *Reverse Auctions: Guidance Is Needed to Maximize Competition and Achieve Cost Savings*. Technical Report. Government Accountability Office. http://www.gao.gov/products/GAO-14-108

Mingyu Guo and Vincent Conitzer. 2010. Computationally Feasible Automated Mechanism Design: General Approach and Case Studies. In *Proceedings of the Twentieth AAAI Conference on Artificial Intelligence (AAAI '10)*. AAAI Press, Atlanta, GA, USA, 1676–1679. http://cs.adelaide.edu.au/~mingyu/publication/nectar2010.pdf NECTAR track.

Jason D. Hartline. 2014. Mechanism Design and Approximation. (Dec. 2014). http://jasonhartline.com/MDnA/ http://jasonhartline.com/MDnA/.

Jason D. Hartline and Tim Roughgarden. 2009. Simple Versus Optimal Mechanisms. In *Proceedings of the 10th ACM Conference on Electronic Commerce (EC '09)*. ACM, New York, NY, USA, 225–234. DOI:http://dx.doi.org/10.1145/1566374.1566407

Interactive Advertising Bureau (IAB). 2015. IAB Programmatic Revenue Report 2014 Results. http://www.iab.net/media/file/PwC_IAB_Programmatic_Study.pdf. (July 2015). http://www.iab.net/media/file/PwC_IAB_Programmatic_Study.pdf

Giuseppe Lopomo. 2001. Optimality and Robustness of the English Auction. *Games and Economic Behavior* 36, 2 (Aug. 2001), 219–240. http://www.sciencedirect.com/science/article/pii/S0899825600908170

D. H. Martin. 1975. On the continuity of the maximum in parametric linear programming. *Journal of Optimization Theory and Applications* 17, 3 (1975), 205–210. DOI:http://dx.doi.org/10.1007/BF00933875

R. Preston McAfee and Philip J. Reny. 1992. Correlated Information and Mechanism Design. *Econometrica* 60, 2 (1992), pp. 395–421. http://www.jstor.org/stable/2951601

Jamie Morgenstern and Tim Roughgarden. 2015. The Pseudo-Dimension of Nearly-Optimal Auctions. In *NIPS*. Forthcoming. http://arxiv.org/abs/1506.03684

Roger B. Myerson. 1981. Optimal Auction Design. *Mathematics of Operations Research* 6, 1 (1981), 58–73. DOI:http://dx.doi.org/10.1287/moor.6.1.58

NASA. 2014. NASA Chooses American Companies to Transport U.S. Astronauts to International Space Station. https://www.nasa.gov/press/2014/september/nasa-chooses-american-companies-to-transport-us-astronauts-to-international. (Sept. 2014). NASAChoosesAmericanCompaniestoTransportU.S.AstronautstoInternationalSpaceStation

Tim Roughgarden and Inbal Talgam-Cohen. 2013. Optimal and Near-optimal Mechanism Design with Interdependent Values. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce (EC '13)*. ACM, New York, NY, USA, 767–784. DOI:http://dx.doi.org/10.1145/2482540.2482606

Tuomas Sandholm and Anton Likhodedov. 2015. Automated Design of Revenue-Maximizing Combinatorial Auctions. *Operations Research* 63, 5 (2015), 1000–1025. DOI:http://dx.doi.org/10.1287/opre.2015.1398