Maximizing Revenue with Limited Correlation: The Cost of Ex-Post Incentive Compatibility

Michael Albert
Department of Computer Science
University of Texas at Austin
2317 Speedway, Stop D9500
Austin, TX 78712, USA
malbert@cs.utexas.edu

Vincent Conitzer
Department of Computer Science
Duke University
Durham, NC 27708, USA
conitzer@cs.duke.edu

Giuseppe Lopomo
Duke University
The Fuqua School of Business
100 Fuqua Drive
Durham, NC 27708, USA
glopomo@duke.edu

Abstract

In a landmark paper in the mechanism design literature, Cremer and McLean (1985) (CM for short) show that when a bidder's valuation is correlated with an external signal, a monopolistic seller is able to extract the full social surplus as revenue. In the original paper and subsequent literature, the focus has been on ex-post incentive compatible (or IC) mechanisms, where truth telling is an ex-post Nash equilibrium. In this paper, we explore the implications of Bayesian versus ex-post IC in a correlated valuation setting. We generalize the full extraction result to settings that do not satisfy the assumptions of CM. In particular, we give necessary and sufficient conditions for full extraction that strictly relax the original conditions given in CM. These more general conditions characterize the situations under which requiring ex-post IC leads to a decrease in expected revenue relative to Bayesian IC. We also demonstrate that the expected revenue from the optimal ex-post IC mechanism guarantees at most a $(|\Theta| + 1)/4$ approximation to that of a Bayesian IC mechanism, where $|\Theta|$ is the number of bidder types. Finally, using techniques from automated mechanism design, we show that, for randomly generated distributions, the average expected revenue achieved by Bayesian IC mechanisms is significantly larger than that for ex-post IC mechanisms.

Introduction

Mechanism design has emerged as a key tool in multi-agent systems for the allocation of tasks and resources to self-interested agents. Though much of the focus of the mechanism design literature has been on independent bidder valuations, many situations, such as the *mineral rights model* (Wilson 1969), are naturally characterized by assuming correlated bidder valuations, so starting with Milgrom and Weber (1982) there has been an increasing body of work on optimal mechanisms in correlated valuation settings. This line of research has led to important, and often quite surprising, insights into many fundamental aspects of mechanism design, the most prominent of which is due to Cremer and McLean (1985). In their landmark paper, they demonstrate that given even small amounts of correlation between bidders' valuations, a monopolistic seller is able to extract the full social surplus as revenue even though the bidders possess private information. This result has become a central focus for the subsequent correlated valuation literature due to its departure from standard results in independent valuation settings. Since the initial Cremer and McLean (1985) paper, much of the subsequent literature has focused on the criticality of the assumptions that lead to full revenue extraction, but one assumption has traditionally been viewed as innocuous: ex-post incentive compatibility (IC for short).

Due to the revelation principle (for a discussion see Krishna (2009)), mechanism design has restricted its attention to incentive compatible mechanisms, i.e. mechanisms where telling the truth about one's type is the optimal strategy for any given agent. However, the specific form of incentive compatibility is dependent on the game theoretic solution concept employed, with the two we will focus on being ex-post IC, corresponding to ex-post Nash equilibrium, and Bayesian IC, corresponding to Bayesian Nash equilibrium, with Bayesian IC being the more permissive. Cremer and McLean (1985) show that under their correlation assumption and ex-interim individual rationality (IR), ex-post IC is sufficient for full revenue extraction, negating the need to examine Bayesian IC separately. While, in follow up work (Cremer and McLean 1988), they are able to relax slightly their initial correlation assumption if they only require Bayesian IC, the literature (McAfee and Reny 1992; Dobzinski, Fu, and Kleinberg 2011; Roughgarden and Talgam-Cohen 2013; Fu et al. 2014) has continued to focus on ex-post IC, due to its sufficiency.

Cremer and McLean (1985; 1988) require extremely precise knowledge by the seller of a very expressive distribution over bidders' beliefs, and when a more realistic assumption is made over the knowledge of the seller, the Cremer and McLean mechanism achieves revenue far from full surplus (Albert, Conitzer, and Lopomo 2015). Therefore, practical mechanism design must account for more limited forms of correlation and less expressive prior distributions. If so, the Cremer and McLean (1985) result is no longer a guarantee for the sufficiency of ex-post IC.

In this paper, we assume ex-interim IR and explore the case of a single bidder whose valuation is correlated with an external signal that both the seller and bidder observe ex-post. This allows us to examine bidder beliefs over the value of the external signal with arbitrary expressiveness, and in so doing, we characterize necessary and sufficient conditions for full surplus extraction in both Bayesian and ex-post IC settings that are a significant relaxation from those of Cremer and McLean (1988). We achieve this by considering the relationship between the valuation function and the information structure. By contrast, Cremer and McLean (1988) characterize the conditions for full revenue extraction solely in terms of the information structure of the problem.

In characterizing our more general conditions, we prove that full revenue extraction is possible for a much larger set of instances under Bayesian IC than ex-post IC. Further, we prove that the expected revenue due an optimal mechanism under ex-post IC achieves at most a $(|\Theta| + 1)/4$ approximation to the optimal Bayesian IC mechanism, where $|\Theta|$ is the number of bidder types. We also use simulation results for randomly generated instances of distributions over bidder types and external signals and thereby demonstrate that Bayesian IC mechanisms generate significantly more revenue than ex-post IC mechanisms. Therefore, our results suggest that for general problems of correlated valuation, it is important to consider Bayesian IC mechanisms.

Our results contribute to recent work on correlated valuations in algorithmic mechanism design (Fu et al. 2014; Roughgarden and Talgam-Cohen 2013; Dobzinski, Fu, and Kleinberg 2011) as well as older work (Lopomo 2001; McAfee and Reny 1992; Milgrom and Weber 1982). By characterizing conditions for full surplus extraction for less expressive distributions we contribute to the literature on learning valuation distributions in mechanism design contexts (Elkind 2007; Blum, Mansour, and Morgenstern 2015; Morgenstern and Roughgarden 2015). We also use methods from automated mechanism design (Conitzer and Sandholm 2002; 2004), and more specifically the use thereof to build intuition for classical results in mechanism design (Likhodedov and Sandholm 2005; Guo and Conitzer 2010). Further, Theorem 3 is reminiscent of the characterization of proper scoring rules (Gneiting and Raftery 2007) and mechanism design as convex analysis (Frongillo and Kash 2014).

Notation and Review of Cremer and McLean

While Cremer and McLean (1988) provides conditions under which a single monopolistic seller selling a single good to some number of buyers n is able to extract full revenue, we restrict our attention to the case of a single monopolistic seller and a single bidder whose valuation is correlated with an external signal, as in Albert, Conitzer, and Lopomo (2015). While our conditions for full surplus extraction can be trivially extended to multiple bidders by considering each in isolation, for the ease and clarity of exposition, we will constrain ourselves to the case of a single bidder. This allows us to focus on the effect of limited correlation, or situations under which the external signal has a smaller state space than the bidder types. This smaller state space for the external signal could be due to the external signal having fewer potential values than the bidder's type, or it could be a consequence of representing many external signals by a reduced set in order to increase the precision of an estimated distribution over the external signals. Additionally, there are many situations under which a single bidder's valuation would be correlated with an external signal, such as a single bidder bidding for unusual keywords for search ads with the external signal being the observed ex-post click through rate for that keyword.

The bidder is risk neutral, and his valuation is drawn from a discrete set of types denoted by $V = \{v_1, v_2, \ldots, v_{|V|}\} \subset [0, \overline{v}]$ where $v_i < v_{i+1}$ for all i. The bidder's type is denoted by $\theta \in \Theta = \{1, 2, \ldots, |\Theta|\}$ and may include information in addition to his valuation. The bidder's valuation is given by $v(\theta)$. We assume without loss of generality that Θ is ordered in increasing valuations so that $v(1) \leq v(2) \leq \ldots \leq v(|\Theta|)$. The discrete external signal is denoted by $\omega \in \Omega = \{1, 2, \ldots, |\Omega|\}$. Let $\Delta(\Omega)$ be the set of all probability measures over Ω. The bidder's type, θ, and the external signal, ω, are drawn from the discrete joint distribution $\boldsymbol{\pi}(\theta, \omega)$ which we will refer to as the *prior*. $\boldsymbol{\pi}(\bullet|\theta) \in \Delta(\Omega)$ is the distribution of the external signal conditional on the bidder type, which we will refer to as the bidder's *belief* over the external signal. Note that $\Delta(\Omega) \subset \mathbb{R}^{|\Omega|}$. We will use bold type to indicate an element $\boldsymbol{f} \in \mathbb{R}^{|\Omega|}$, with $\boldsymbol{f}(i)$ denoting the ith element of the vector $\boldsymbol{f}$.

A (direct revelation) mechanism is defined by, given the bidder's type and the observed external signal, the probability that the seller allocates the item to the bidder and a transfer from the bidder to the seller. We will denote the probability of allocation by $q(\theta, \omega)$, which must be between zero and one, and the transfer by $m(\theta, \omega)$. Note that the transfer $m(\theta, \omega)$ can be either positive or negative with a positive amount indicating a transfer from the bidder to the seller and a negative amount a transfer from the seller to the bidder.

Definition 1 (Bidder's Utility)**.** *Given true type $\theta \in \Theta$, reported type $\hat{\theta} \in \Theta$, and external signal $\omega \in \Omega$, the bidder's utility under mechanism (q, m) is:*

$$U(\theta, \hat{\theta}, \omega) = v(\theta)q(\hat{\theta}, \omega) - m(\hat{\theta}, \omega) \tag{1}$$

The bidder's expected utility is given by:

$$U(\theta, \hat{\theta}) = \sum_{\omega} (v(\theta)q(\hat{\theta}, \omega) - m(\hat{\theta}, \omega))\boldsymbol{\pi}(\omega|\theta) \tag{2}$$

Since the bidder has a private valuation, the seller must induce the bidder to reveal his private information. He does this by ensuring that the bidder always finds it optimal, or incentive compatible, to truthfully reveal his private information.

Definition 2 (Ex-Post Incentive Compatibility)**.** *A mechanism (q, m) is* ex-post incentive compatible (IC) *if:*

$$\forall \theta, \hat{\theta} \in \Theta, \omega \in \Omega : U(\theta, \theta, \omega) \geq U(\theta, \hat{\theta}, \omega) \tag{3}$$

Definition 3 (Bayesian Incentive Compatibility)**.** *A mechanism (q, m) is* Bayesian incentive compatible (IC) *if:*

$$\forall \theta, \hat{\theta} \in \Theta : U(\theta, \theta) \geq U(\theta, \hat{\theta}) \tag{4}$$

Note that ex-post IC is more strict than Bayesian IC in the sense that every ex-post IC mechanism is also Bayesian IC. In addition to IC, the bidder will not participate in the mechanism if he expects to be made worse off by participating. Therefore, we require that in expectation, the bidder will receive nonnegative utility.

Definition 4 (Ex-Interim Individual Rationality). *A mechanism* (q, m) *is* ex-interim individually rational (IR) *if:*

$$\forall \theta \in \Theta : U(\theta, \theta) \geq 0 \tag{5}$$

It is important to note that ex-interim IR is an essential assumption for the Cremer and McLean (1985) full surplus extraction result. Ex-interim IR only ensures that the bidder has non-negative utility in expectation, but it allows for negative utility under certain realizations of the external signal. Lopomo (2001) shows that if this condition is replaced with one that guarantees non-negative utility under all realizations of the external signal, full surplus extraction is no longer possible in the general case.

Further, ex-interim IC has been criticized as not being a robust assumption, as in being very sensitive to the specification of the prior (Roughgarden and Talgam-Cohen 2013). However, given its central importance to the Cremer and McLean (1985) result, it is widely used in the literature (Dobzinski, Fu, and Kleinberg 2011; Fu et al. 2014; Albert, Conitzer, and Lopomo 2015).

The common argument against the use of Bayesian IC (Roughgarden and Talgam-Cohen 2013) is that it generates mechanisms that are sensitive to the specification of the prior. However, given the use of ex-interim IR, the additional sensitivity imparted by Bayesian IC is unlikely to be a first order effect.

Definition 5 (Optimal Mechanisms). *A mechanism* (q, m) *is an* optimal ex-post mechanism *if under the constraint of ex-interim individual rationality and ex-post incentive compatibility it maximizes the following:*

$$\sum_{\theta,\omega} m(\theta, \omega)\boldsymbol{\pi}(\theta, \omega) \tag{6}$$

A mechanism that maximizes the above under the constraint of ex-interim individual rationality and Bayesian incentive compatibility is an optimal Bayesian mechanism*.*

Definition 6 (Full Social Surplus Extraction as Revenue). *We say that a mechanism* extracts the full social surplus as revenue in expectation *if there exists a (Bayesian or ex-post) mechanism such that:*

$$\sum_{\theta,\omega} \boldsymbol{\pi}(\theta, \omega)m(\theta, \omega) = \sum_{\theta,\omega} \boldsymbol{\pi}(\theta, \omega)v(\theta). \tag{7}$$

Cremer and McLean (1985) are able to extract full social surplus by combining a VCG mechanism with a lottery over the outcome of the external signal. However, in order to ensure that they can construct the lottery in a way that is incentive compatible and individually rational, they must make the following assumption concerning bidder beliefs over the external signal.

Assumption 1. *For all* $\theta \in \Theta$*, let* $\boldsymbol{\Gamma}$ *be the following matrix whose rows are indexed by the* $|\Omega|$ *elements of* Ω*, and whose columns are indexed by the* $|\Theta|$ *elements of* Θ*:*

$$\boldsymbol{\Gamma} = \begin{bmatrix} \boldsymbol{\pi}(1|1) & \cdots & \boldsymbol{\pi}(|\Omega||1) \\ \vdots & \ddots & \vdots \\ \boldsymbol{\pi}(1||\Theta|) & \cdots & \boldsymbol{\pi}(|\Omega|||\Theta|) \end{bmatrix} \tag{8}$$

$\boldsymbol{\Gamma}$ *has rank* $|\Theta|$*.*

Theorem 1 (Cremer and McLean (1985)). *Under Assumption 1, there exists an ex-post IC mechanism that extracts the full social surplus as revenue in expectation.*

A proof of Theorem 1 can be found in Cremer and McLean (1985) or Krishna (2009).

One of the implications of Assumption 1 is that $|\Omega| \geq |\Theta|$. While this assumption may seem reasonable, particularly if the external signal is another bidder's valuation as in a multi-bidder setting, this requires very precise knowledge of the bidder's beliefs. As was shown by Albert, Conitzer and Lopomo (2015), even slight uncertainties over the bidder's beliefs can cause the optimal mechanism to perform equivalently to simple mechanisms, such as take it or leave it offers. Given the necessity of a precise estimation of the conditional distribution, it may be necessary to combine multiple values for the external signal into a single value in order to create a more precisely estimated, but less expressive, prior distribution.

Finally, we will make use of the notion of a subgradient.

Definition 7. *Given some function* $G : \mathbb{R}^{|\Omega|} \to \mathbb{R}$ *a* subgradient to G at $\boldsymbol{f} \in \mathbb{R}^{|\Omega|}$ *is a linear function* $d : \mathbb{R}^{|\Omega|} \to \mathbb{R}$ *such that* $d(\mathbf{0}) = 0$ *and for all* $\boldsymbol{g} \in \mathbb{R}^{|\Omega|}$*,*

$$G(\boldsymbol{g}) \geq G(\boldsymbol{f}) + d(\boldsymbol{g} - \boldsymbol{f}). \tag{9}$$

Necessary and Sufficient Conditions for Full Surplus Extraction as Revenue

In this section, we characterize necessary and sufficient conditions for full social surplus extraction as revenue given arbitrary correlation structures. Our results guarantee full surplus extraction even when $|\Omega| < |\Theta|$ or a subset of the conditional beliefs are a linear combination of others, both of which violate Assumption 1. We are able to extend Cremer and McLean (1985)'s result by considering the interaction between the prior $\boldsymbol{\pi}$ and the valuation function. We will use the following lemma.

Lemma 1. *A mechanism* (q, m) *extracts full surplus if and only if* $q(\theta, \omega) = 1$ *and* $U(\theta, \theta) = 0$ *for all* $\theta \in \Theta, \omega \in \Omega$*.*

The proof of Lemma 1 is straightforward.

Theorem 2 (Full Surplus Extraction with Ex-Post IC). *For a given* $(\boldsymbol{\pi}, V, \Omega)$*, full surplus extraction is possible for an ex-post incentive compatible mechanism if and only if there exists a* linear *function* $G : \mathbb{R}^{|\Omega|} \to \mathbb{R}$ *such that* $G(\boldsymbol{\pi}(\bullet|\theta)) = -v(\theta)$*.*

Proof. First, assume that there exists an ex-post incentive compatible mechanism (q, ω) such that full surplus extraction is achieved in expectation. Then, by Definition 2 and Lemma 1:

$$\forall \theta, \hat{\theta} \in \Theta, \omega \in \Omega : v(\theta) - m(\theta, \omega) = v(\theta) - m(\hat{\theta}, \omega).$$

Therefore,

$$\forall \theta, \hat{\theta} \in \Theta, \omega \in \Omega : m(\theta, \omega) = m(\hat{\theta}, \omega) = m^*(\omega).$$

Set, for $\boldsymbol{f} \in \mathbb{R}^{|\Omega|}$, $G(\boldsymbol{f}) = \sum_{\omega\in\Omega} m^*(\omega)\boldsymbol{f}(\omega)$. Then, for $a, b \in \mathbb{R}$ and $\boldsymbol{f}, \boldsymbol{g} \in \mathbb{R}^{|\Omega|}$,

$$\begin{aligned}G(a\boldsymbol{f} + b\boldsymbol{g}) &= \sum_{\omega\in\Omega} m^*(\omega)(a\boldsymbol{f}(\omega) + b\boldsymbol{g}(\omega)) \\ &= a\sum_{\omega\in\Omega} m^*(\omega)\boldsymbol{f}(\omega) + b\sum_{\omega\in\Omega} m^*(\omega)\boldsymbol{g}(\omega) \\ &= aG(\boldsymbol{f}) + bG(\boldsymbol{g})\end{aligned}$$

Further, $G(\boldsymbol{\pi}(\bullet|\theta)) = \sum_{\omega\in\Omega} m(\theta,\omega)f(\omega) = -v(\theta)$ by ex-interim IR and Lemma 1.

Alternatively, suppose that there does exist a linear function $G : \mathbb{R}^{|\Omega|} \to \mathbb{R}$ such that $G(\boldsymbol{\pi}(\bullet|\theta)) = -v(\theta)$. Denote by $\boldsymbol{\pi}_i^* \in \Delta(\Omega)$ the probability distribution such that $\boldsymbol{\pi}_i^*(\omega = i) = 1$ and $\boldsymbol{\pi}_i^*(\omega \neq i) = 0$. For all $\theta \in \Theta$ and $\omega \in \Omega$, set $q(\theta,\omega) = 1$ and $m(\theta,\omega) = G(\boldsymbol{\pi}_\omega^*)$. Then,

$$\begin{aligned}\sum_{\omega\in\Omega} m(\theta,\omega)\boldsymbol{\pi}(\omega|\theta) &= \sum_{\omega\in\Omega} G(\boldsymbol{\pi}_\omega^*)\boldsymbol{\pi}(\omega|\theta)) \\ &= G\left(\sum_{\omega\in\Omega} \boldsymbol{\pi}_\omega^*\boldsymbol{\pi}(\omega|\theta)\right) \\ &= G(\boldsymbol{\pi}(\bullet|\theta)) = -v(\theta).\end{aligned}$$

Therefore, ex-interim IR is binding, and given that for all $\theta, \hat{\theta} \in \Theta$, $m(\theta,\omega) = m(\hat{\theta},\omega) = m^*(\omega) = G(\boldsymbol{\pi}_i^*)$, ex-post IC is satisfied. □

Intuitively, Theorem 2 states that ex-post IC combined with ex-interim IR allow for the mechanism designer to incorporate a single lottery over the external signal into the ex-post mechanism. This lottery is such that the payoff for $\omega = i$ is the linear function G evaluated at $\boldsymbol{\pi}_i^*$ (defined as in the above proof). Full surplus extraction is only possible, then, when one lottery, or linear function, can intersect every valuation. By contrast, the additional power of a Bayesian mechanism is that the mechanism designer can incorporate many lotteries.

Theorem 3 (Full Surplus Extraction with Bayesian IC)**.** *For a given $(\boldsymbol{\pi}, V, \Omega)$, full surplus extraction is possible for a Bayesian incentive compatible mechanism if and only if there exists a* convex *function $G : \mathbb{R}^{|\Omega|} \to \mathbb{R}$ such that $G(\boldsymbol{\pi}(\bullet|\theta)) = -v(\theta)$.*

Proof. Assume that there exists a Bayesian incentive compatibility mechanism (q, m) such that the full surplus is extracted in expectation. Suppose in addition that there does not exist a convex function $G : \mathbb{R}^{|\Omega|} \to \mathbb{R}$ such that $G(\boldsymbol{\pi}(\bullet|\theta)) = -v(\theta)$. This implies, by the definition of convexity, that for every function $G^* : \mathbb{R}^{|\Omega|} \to \mathbb{R}$ such that $G^*(\boldsymbol{\pi}(\bullet|\theta)) = -v(\theta)$ there must exist $\theta^* \in \Theta$ such that for $\theta \in \Theta\backslash\{\theta^*\}$, there exists $\alpha_\theta \geq 0$ where $\sum_{\theta\in\Theta\backslash\{\theta^*\}} \alpha_\theta = 1$, $\boldsymbol{\pi}(\omega|\theta^*) = \sum_{\theta\in\Theta\backslash\{\theta^*\}} \alpha_\theta\boldsymbol{\pi}(\omega|\theta)$, and $\sum_{\theta\in\Theta\backslash\{\theta^*\}} \alpha_\theta(-v(\theta)) = \sum_{\theta\in\Theta\backslash\{\theta^*\}} \alpha_\theta G^*(\boldsymbol{\pi}(\bullet|\theta)) < G^*(\sum_{\theta\in\Theta\backslash\{\theta^*\}} \alpha_\theta\boldsymbol{\pi}(\omega|\theta)) = G^*(\boldsymbol{\pi}(\bullet|\theta^*)) = -v(\theta^*)$.

Note that for all $\theta \in \Theta$, $U(\theta,\theta) = 0$ by Lemma 1. Then

$$\begin{aligned}&\sum_{\theta\in\Theta\backslash\{\theta^*\}} \alpha_\theta U(\theta,\theta^*) \\ &\quad= \sum_{\theta\in\Theta\backslash\{\theta^*\}} \alpha_\theta(\sum_\omega(v(\theta) - m(\theta^*,\omega))\boldsymbol{\pi}(\omega|\theta) \\ &\quad> v(\theta^*) - \sum_\omega m(\theta^*,\omega) \sum_{\theta\in\Theta\backslash\{\theta^*\}} \alpha_\theta\boldsymbol{\pi}(\omega|\theta) \\ &\quad= v(\theta^*) - \sum_\omega m(\theta^*,\omega)\boldsymbol{\pi}(\omega|\theta^*) \\ &\quad= v(\theta^*) - v(\theta^*) = 0\end{aligned}$$

Then, there exists $\theta' \in \Theta\backslash\{\theta^*\}$ such that $U(\theta',\theta^*) > 0 = U(\theta',\theta')$, in contradiction of Bayesian IC. Therefore, there must exist a convex function $G : \mathbb{R}^{|\Omega|} \to \mathbb{R}$ such that $G(\boldsymbol{\pi}(\bullet|\theta)) = -v(\theta)$.

Now, assume that there does exist a convex function $G : \mathbb{R}^{|\Omega|} \to \mathbb{R}$ such that $G(\boldsymbol{\pi}(\bullet|\theta)) = -v(\theta)$. Let d_θ be a subgradient to G at $\boldsymbol{\pi}(\bullet|\theta)$. This exists everywhere by the convexity of G. Denote by $\boldsymbol{\pi}_i^* \in \Delta(\Omega)$ the probability distribution such that $\boldsymbol{\pi}_i^*(\omega = i) = 1$ and $\boldsymbol{\pi}_i^*(\omega \neq i) = 0$. Then, for all $\theta \in \Theta$ and $\omega \in \Omega$ set $q(\theta,\omega) = 1$, and the transfers such that

$$m(\theta,\omega) = -(G(\boldsymbol{\pi}(\bullet|\theta)) + d_\theta(\boldsymbol{\pi}_\omega^* - \boldsymbol{\pi}(\bullet|\theta))).$$

Then

$$\begin{aligned}U(\theta,\theta) &= \sum_\omega(v(\theta) - m(\theta,\omega))\boldsymbol{\pi}(\omega|\theta) \\ &= v(\theta) + \sum_\omega(G(\boldsymbol{\pi}(\bullet|\theta)) + d_\theta(\boldsymbol{\pi}_\omega^* - \boldsymbol{\pi}(\bullet|\theta)))\boldsymbol{\pi}(\omega|\theta) \\ &= v(\theta) + G(\boldsymbol{\pi}(\bullet|\theta)) + d_\theta\left(\sum_\omega \boldsymbol{\pi}_\omega^*\boldsymbol{\pi}(\omega|\theta) - \boldsymbol{\pi}(\bullet|\theta)\right) \\ &= v(\theta) - v(\theta) + d(0) = 0.\end{aligned}$$

Therefore, ex-interim IR binds for all $\theta \in \Theta$. Also, for all $\theta, \hat{\theta} \in \Theta$

$$\begin{aligned}U(\theta,\hat{\theta}) &= \sum_\omega(v(\theta) - m(\hat{\theta},\omega))\boldsymbol{\pi}(\omega|\theta) \\ &= v(\theta) + \sum_\omega(G(\boldsymbol{\pi}(\bullet|\hat{\theta})) + d_\theta(\boldsymbol{\pi}_\omega^* - \boldsymbol{\pi}(\bullet|\hat{\theta})))\boldsymbol{\pi}(\omega|\theta) \\ &= v(\theta) + G(\boldsymbol{\pi}(\bullet|\hat{\theta})) + d_\theta(\sum_\omega \boldsymbol{\pi}_\omega^*\boldsymbol{\pi}(\omega|\theta) - \boldsymbol{\pi}(\bullet|\hat{\theta})) \\ &= -G(\boldsymbol{\pi}(\bullet|\theta)) + G(\boldsymbol{\pi}(\bullet|\hat{\theta})) + d_\theta(\boldsymbol{\pi}(\bullet|\theta) - \boldsymbol{\pi}(\bullet|\hat{\theta})) \\ &\leq 0 = U(\theta,\theta)\end{aligned}$$

Therefore, by Lemma 1, the mechanism extracts full surplus. □

Theorem 3 is able to relax the necessity of a linear function G by using multiple lotteries. Each lottery corresponds to a linear function, just as in Theorem 2, but the linear function is a subgradient of a convex function. The convexity of the function ensures that each bidder finds it IC to only participate in the lottery that corresponds to his type. Figure 1

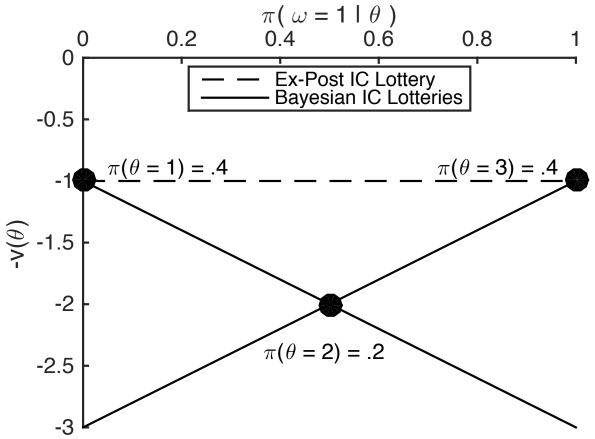

Figure 1: The optimal ex-post mechanism always allocates the item to all three types and sets for all $\theta \in \Theta$, $m(\theta, 1) = m(\theta, 2) = 1$, a lottery with a constant payoff, giving revenue of 1. The optimal Bayesian mechanism extracts full revenue by always allocating the item and setting $m(1, 2) = m(2, 2) = 1$, $m(1, 1) = m(2, 1) = 3$ (the same lottery for $\theta = 1$ and $\theta = 2$), $m(3, 1) = 3$, and $m(3, 2) = 1$, generating revenue of 1.2.

depicts this graphically for $|\Theta| = 3$ and $|\Omega| = 2$. Note that this result is reminiscent of *proper scoring rules* (Gneiting and Raftery 2007; Frongillo and Kash 2014).

Corollary 2. *Assumption 1 implies that there exists a linear function $G : \mathbb{R}^{|\Omega|} \to \mathbb{R}$ such that $G(\boldsymbol{\pi}(\bullet|\theta)) = -v(\theta)$.*

Proof. Let Γ be defined as in Assumption 1, and let $\boldsymbol{v} \in \mathbb{R}^{|\Theta|}$ be such that $\boldsymbol{v}(\theta) = v(\theta)$. Then by the assumption of full rank, there exists $\boldsymbol{c} \in \mathbb{R}^{|\Omega|}$ such that $\Gamma \boldsymbol{c} = -\boldsymbol{v}$.

Define $G : \mathbb{R}^{|\Omega|} \to \mathbb{R}$ such that for $\boldsymbol{f} \in \mathbb{R}^{|\Omega|}$, $G(\boldsymbol{f}) = \sum_{\omega \in \Omega} \boldsymbol{c}(\omega)\boldsymbol{f}(\omega)$. Then, G is linear and $G(\boldsymbol{\pi}(\bullet|\theta)) = -v(\theta)$. □

Theorem 4. *The expected revenue generated by an optimal ex-post mechanism guarantees at most a $(|\Theta| + 1)/4$ approximation to the expected revenue generated by an optimal Bayesian mechanism*

Proof. Let $\Theta = \{1, 2, ..., |\Theta|\}$ and $v(\theta) = 2^\theta$ for $\theta \in \Theta$. Let $\boldsymbol{\pi}(\theta) = \sum_\omega \boldsymbol{\pi}(\theta, \omega) = 1/2^\theta$ except for $\boldsymbol{\pi}(|\Theta|) = 1/2^{|\Theta|-1}$. $|\Omega| = 2$, and $\boldsymbol{\pi}(1|\theta) = (2^{k(\theta-1)} - 1)/(2^{k(|\Theta|-1)} - 1)$, where $k > 1$. Then, if $G(\boldsymbol{f}) = -2((2^{k(|\Theta|-1)} - 1)\boldsymbol{f}(1) + 1)^{-1/k}$, $G(\boldsymbol{\pi}(\bullet|\theta)) = -v(\theta)$. Note that $G''(\boldsymbol{f}) > 0$ for $k > 1$, so G is convex.

Since G is convex, the Bayesian IC mechanism extracts full surplus, and full surplus is $|\Theta| + 1$. Also, there does not exist a linear function H such that $H(\boldsymbol{\pi}(\bullet|\theta)) = -v(\theta)$, so the optimal ex-post IC mechanism will not extract full surplus.

Note that for an ex-post IC mechanism, given ω, it can be shown by direct application of ex-post IC that for all $\theta \in \Theta$, $q(\theta + 1, \omega) \geq q(\theta, \omega)$. Further, for all $\theta \in \Theta, \omega \in \Omega$ such that $q(\theta, \omega) = 1$, $m(\theta, \omega) = m(\omega)$, again by direct application of ex-post IC. Finally, for any optimal ex-post IC mechanism, $q(|\Theta|, \omega) = 1$ for all $\omega \in \Omega$.

Let (q^*, m^*) be an optimal ex-post IC mechanism. Suppose that for $\omega = 2$, there exists a $\theta' \in \Theta$ such that $q^*(\theta', 2) < 1$ and $q^*(\theta' + 1, 2) = 1$. Also, by ex-post IC and the assumed optimatiliy, $v(\theta' + 1) - m^*(2) = v(\theta' + 1)q^*(\theta', 2) - m^*(\theta', 2)$, because if not $q^*(\theta, 2)$ and $m^*(\theta', 2)$ could be increased, increasing expected revenue and violating the optimality of the mechanism. Therefore, $m^*(2) = v(\theta' + 1)(1 - q^*(\theta', 2)) + m^*(\theta', 2)$ Define a new mechanism (q', m') such that $q'(\theta', 2) = 1$, $m'(\theta', 2) = m^*(\theta', 2) + v(\theta')(1 - q^*(\theta', 2))$, $m'(2) = m'(\theta', 2)$, and for all other (θ, ω), $q'(\theta, \omega) = q^*(\theta, \omega)$ and $m'(\theta, \omega) = m^*(\theta, \omega)$. Therefore, $m'(\theta', 2) - m^*(\theta, 2) = v(\theta')(1 - q(\theta', 2))$ and $m'(2) - m^*(2) = (v(\theta') - v(\theta' + 1))(1 - q(\theta', 2))$. Note that (q', m') is ex-post IC and ex-interim IR.

Then, the difference in expected revenue between mechanism (q', m') and (q^*, m^*) is

$$
\begin{aligned}
&\sum_{\theta,\omega} \boldsymbol{\pi}(\theta, \omega)m'(\theta, \omega) - \sum_{\theta,\omega} \boldsymbol{\pi}(\theta, \omega)m^*(\theta, \omega) \\
&= \sum_{\theta=\theta'+1}^{|\Theta|} \frac{1}{2^\theta}\boldsymbol{\pi}(2|\theta)(m'(2) - m^*(2)) \\
&\quad + \frac{1}{2^{\theta'}}\boldsymbol{\pi}(2|\theta')(m'(\theta', 2) - m^*(\theta', 2)) \\
&= -\sum_{\theta=\theta'+1}^{|\Theta|} \frac{1}{2^\theta}\boldsymbol{\pi}(2|\theta)2^{\theta'}(1 - q^*(\theta', 2)) \\
&\quad + \frac{1}{2^{\theta'}}\boldsymbol{\pi}(2|\theta')2^{\theta'}(1 - q^*(\theta', 2)) > 0
\end{aligned}
$$

Therefore, (q^*, m^*) is not optimal, which is a contradiction. This implies that for an optimal ex-post IC mechanism (q^*, m^*) for all $\theta \in \Theta$, $q^*(\theta, 2) = 1$, and by ex-interim IR applied at $\theta = 1$, $m^*(2) = 2$. Also, since for all $\theta' \in \Theta\backslash\{|\Theta|\}$, $\pi(|\Theta|, 1)v(|\Theta|) = 2 > \sum_{\theta=1}^{|\Theta|-1} \pi(\theta, 1)v(\theta')$, ex-post IR must bind for $\theta = |\Theta|$, which implies $m^*(|\Theta|, 1) = 2^{|\Theta|}$. Further, it is trivial to verify that for all $\theta \in \Theta$, $q^*(\theta, 1) = 1$ and $m^*(\theta, 1) = m^*(|\Theta|, 1)$ is ex-post IC and ex-interim IR. Therefore, the mechanism (q, m) where for all $\theta \in \Theta$ and $\omega \in \Omega$, $q(\theta, \omega) = 1$, $m(1) = 2^{|\Theta|}$, and $m(2) = 2$ is an optimal ex-post IC mechanism.

For sufficiently large k, it is easy to verify that $\sum_{\theta \in \Theta\backslash\{|\Theta|\}}(1/2^\theta(2^{|\Theta|}\boldsymbol{\pi}(1|\theta) + 2\boldsymbol{\pi}(2|\theta))) < 2$. Choose k such that this is true. Then, the expected revenue due to an optimal ex-post IC mechanism is given by:

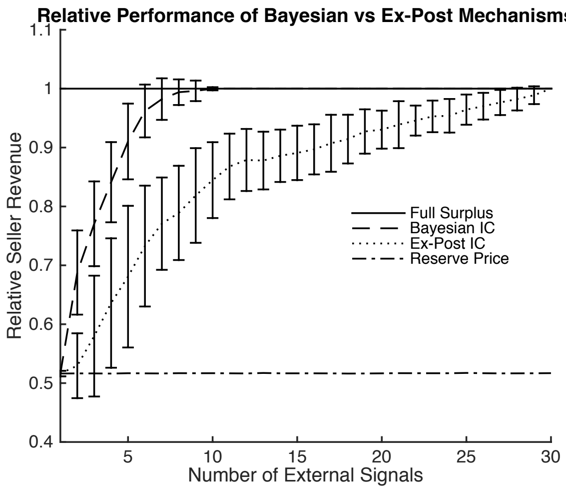

Figure 2: Relative performance of Bayesian vs Ex-Post IC mechanisms for randomly generated distributions. $|\Theta| = 30$, $\rho = .1$, and we vary $|\Omega|$ from $\{1, ..., 30\}$. Each point is the average of 100 randomly generated distributions. 95% confidence intervals shown.

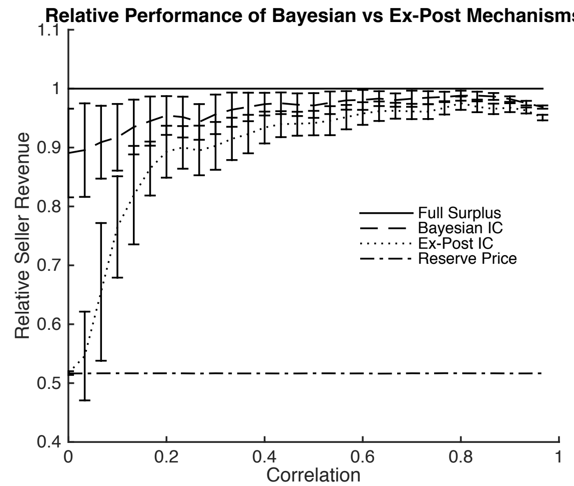

Figure 3: Relative performance of Bayesian vs Ex-post IC mechanisms for randomly generated distributions. $|\Theta| = 30$, $|\Omega| = 5$, and we vary ρ from $[0, 1]$ with a step size of $1/30$. Each point is the average of 100 randomly generated distributions. 95% confidence intervals shown.

$$\begin{aligned}\sum_{\theta,\omega} &\boldsymbol{\pi}(\theta,\omega)m(\theta,\omega) \\ &= \sum_{\theta\in\Theta\setminus\{|\Theta|\}} \frac{1}{2^{\theta}}\left(2^{|\Theta|}\boldsymbol{\pi}(1|\theta) + 2\boldsymbol{\pi}(2|\theta)\right) \\ &\quad + (1/2^{|\Theta|-1})2^{|\Theta|} < 4.\end{aligned}$$

Therefore, the optimal Bayesian IC mechanism has revenue of $|\Theta| + 1$ and the optimal ex-post IC mechanism has revenue less than 4. □

Simulation Results

To explore the relative importance of Bayesian versus ex-post IC for revenue efficiency, we generate random distributions and solve for the optimal mechanism using automated mechanism design techniques (Conitzer and Sandholm 2002; Albert, Conitzer, and Lopomo 2015) under both instances of Bayesian and ex-post IC. We construct each distribution by generating 1500 samples, where a sample consists of two independent draws from the uniform distribution over $[0, 1]$ corresponding to a realization of θ and ω. Denote sample i by $\boldsymbol{x}_i = (\theta_i, \omega_i)$. We pick a target correlation ρ, and construct the 2×2 correlation matrix C. Then, we decompose the correlation matrix using the Cholesky decomposition to calculate L such that $C = LL^T$. Finally, we transform our independent values for θ and ω into correlated values by constructing the final sample as $\boldsymbol{x}'_i = \boldsymbol{x}_i L^T$. This guarantees that $corr(\boldsymbol{x}'(1), \boldsymbol{x}'(2)) = \rho$ within the sample.

We then construct a discrete probability distribution from the 1500 samples by calculating equally spaced buckets along $[0, 1]^2$, where the number of buckets along each dimension is equal to the number of bidder types and number of external signals. Then we count how many samples fall in each bucket and normalize. The valuations corresponding to each sample is the upper limit of the bucket that contains the sample. Note that due to the way that we correlate the two independent variables, there are some samples where $\omega_i > 1$. In that case, we count them in the highest bucket.

This procedure allows us to generate correlated joint distributions randomly with each distribution unique, as well as allowing us to independently vary both the number of external signals and the degree of correlation. Note that with probability 1, if the number of external signals $|\Omega|$ is equal to the number of valuations $|\Theta|$, then Assumption 1 will be satisfied and full social surplus extraction will be possible as in Cremer and McLean (1985). This guarantees that as we increase the number of external signals, we should converge to full social surplus extraction. However, as can be seen in Figure 2, the optimal Bayesian mechanism converges to full extraction with much fewer external signals than the optimal ex-post mechanism, becoming indistinguishable from full surplus extraction with $|\Omega| = 10$, while the optimal ex-post IC mechanism generates significantly less revenue until the number of external states equals the number of bidder types.

In Figure 3, we vary the degree of correlation, while holding the number of external states constant. We observe that as the correlation between the bidder's type and the external signal approaches 1, both the optimal Bayesian and ex-post mechanisms get very close to full revenue extraction. However, the optimal Bayesian mechanism generates statistically significant higher revenue than the optimal ex-post mechanism for all correlation values.

Conclusion

Due to the sufficiency of ex-post IC for full surplus extraction in Cremer and McLean (1985), it is widely used in optimal mechanism design with correlated valuations. However, it is crucial to understand the effect of our mechanism design choices when the assumptions of Cremer and McLean (1985) fail, which they are likely to do in practice. We generalize their result to settings of more limited correlation, proving conditions for full surplus extraction, and in the process we demonstrate both theoretically and empirically that Bayesian IC is important for maximizing revenue under our less restrictive assumptions. Our empirical results suggest that Bayesian IC is likely to be particularly important for settings where the external signal state space is small (either intrinisically or due to distribution estimation concerns) or when correlation is low. Given our results, we believe that any attempt to implement an optimal mechanism by learning the beliefs of bidders over external signals would likely benefit from focusing on Bayesian IC mechanisms.

Acknowledgements

Conitzer is thankful for support from NSF under awards IIS-1527434, IIS-0953756, CCF-1101659, and CCF-1337215, ARO under grants W911NF-12-1-0550 and W911NF-11-1-0332, and a Guggenheim Fellowship. This work was done in part while Conitzer was visiting the Simons Institute for the Theory of Computing.

We would also like to thank Peter Stone for very helpful comments during the course of this work.

References

Albert, M.; Conitzer, V.; and Lopomo, G. 2015. Assessing the robustness of Cremer-Mclean with automated mechanism design. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, AAAI '15, 763–769. Austin, TX, USA: AAAI Press.

Blum, A.; Mansour, Y.; and Morgenstern, J. 2015. Learning valuation distributions from partial observation. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, AAAI '15, 798–805. Austin, TX, USA: AAAI Press.

Conitzer, V., and Sandholm, T. 2002. Complexity of mechanism design. In *Proceedings of the 18th Annual Conference on Uncertainty in Artificial Intelligence*, UAI '02, 103–110. Alberta, Canada: Morgan Kaufmann.

Conitzer, V., and Sandholm, T. 2004. Self-interested automated mechanism design and implications for optimal combinatorial auctions. In *Proceedings of the 5th ACM Conference on Electronic Commerce*, EC '04, 132–141. New York, NY, USA: ACM.

Cremer, J., and McLean, R. P. 1985. Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica* 53(2):pp. 345–361.

Cremer, J., and McLean, R. P. 1988. Full extraction of the surplus in bayesian and dominant strategy auctions. *Econometrica* 56(6):pp. 1247–1257.

Dobzinski, S.; Fu, H.; and Kleinberg, R. D. 2011. Optimal auctions with correlated bidders are easy. In *Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing*, STOC '11, 129–138. New York, NY, USA: ACM.

Elkind, E. 2007. Designing and learning optimal finite support auctions. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '07, 736–745. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics.

Frongillo, R., and Kash, I. 2014. General truthfulness characterizations via convex analysis. In *Web and Internet Economics*. Springer. 354–370.

Fu, H.; Haghpanah, N.; Hartline, J.; and Kleinberg, R. 2014. Optimal auctions for correlated buyers with sampling. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation*, EC '14, 23–36. New York, NY, USA: ACM.

Gneiting, T., and Raftery, A. E. 2007. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association* 102(477):359–378.

Guo, M., and Conitzer, V. 2010. Computationally feasible automated mechanism design: General approach and case studies. In *Proceedings of the Twentieth AAAI Conference on Artificial Intelligence*, AAAI '10, 1676–1679. Atlanta, GA, USA: AAAI Press. NECTAR track.

Krishna, V. 2009. *Auction theory*. Academic press, 2nd edition.

Likhodedov, A., and Sandholm, T. 2005. Approximating revenue-maximizing combinatorial auctions. In *Proceedings of the Nineteenth AAAI Conference on Artificial Intelligence*, AAAI '05, 267–274. Pittsburgh, PA, USA: AAAI Press.

Lopomo, G. 2001. Optimality and robustness of the english auction. *Games and Economic Behavior* 36(2):219–240.

McAfee, R. P., and Reny, P. J. 1992. Correlated information and mechanism design. *Econometrica* 60(2):pp. 395–421.

Milgrom, P. R., and Weber, R. J. 1982. A theory of auctions and competitive bidding. *Econometrica* 50(5):pp. 1089–1122.

Morgenstern, J., and Roughgarden, T. 2015. The pseudo-dimension of nearly-optimal auctions. In *NIPS*, Forthcoming.

Roughgarden, T., and Talgam-Cohen, I. 2013. Optimal and near-optimal mechanism design with interdependent values. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, EC '13, 767–784. New York, NY, USA: ACM.

Wilson, R. B. 1969. Competitive bidding with disparate information. *Management Science* 15(7):446–452.