Empirical Mechanism Design: Methods, with Application to a Supply-Chain Scenario

Yevgeniy Vorobeychik, Christopher Kiekintveld, and Michael P. Wellman
University of Michigan
Computer Science & Engineering
Ann Arbor, MI 48109-2121 USA
{ yvorobey, ckiekint, wellman }@umich.edu

ABSTRACT

Our proposed methods employ learning and search techniques to estimate outcome features of interest as a function of mechanism parameter settings. We illustrate our approach with a design task from a supply-chain trading competition. Designers adopted several rule changes in order to deter particular procurement behavior, but the measures proved insufficient. Our empirical mechanism analysis models the relation between a key design parameter and outcomes, confirming the observed behavior and indicating that no reasonable parameter settings would have been likely to achieve the desired effect. More generally, we show that under certain conditions, the estimator of optimal mechanism parameter setting based on empirical data is consistent.

Categories and Subject Descriptors

I.6 [**Computing Methodologies**]: Simulation and Modeling; J.4 [**Computer Applications**]: Social and Behavioral Sciences—*Economics*

General Terms

Algorithms, Economics, Design

Keywords

Empirical Mechanism Design, Game Theory

1. MOTIVATION

We illustrate our problem with an anecdote from a supply chain research exercise: the 2003 and 2004 Trading Agent Competition (TAC) Supply Chain Management (SCM) game. TAC/SCM [1] defines a scenario where agents compete to maximize their profits as manufacturers in a supply chain. The agents procure components from the various suppliers and assemble finished goods for sale to customers, repeatedly over a simulated year.[1]

As it happened, the specified negotiation behavior of suppliers provided a great incentive for agents to procure large quantities of components on day 0: the very beginning of the simulation. During the early rounds of the 2003 SCM competition, several agent developers discovered this, and the apparent success led to most agents performing the majority of their purchasing on day 0. Although jockeying for day-0 procurement turned out to be an interesting strategic issue in itself [19], the phenomenon detracted from other interesting problems, such as adapting production levels to varying demand (since component costs were already sunk), and dynamic management of production, sales, and inventory. Several participants noted that the predominance of day-0 procurement overshadowed other key research issues, such as factory scheduling [2] and optimizing bids for customer orders [13]. After the 2003 tournament, there was a general consensus in the TAC community that the rules should be changed to deter large day-0 procurement.

The task facing game organizers can be viewed as a problem in *mechanism design*. The designers have certain game features under their control, and a set of objectives regarding game outcomes. Unlike most academic treatments of mechanism design, the objective is a behavioral feature (moderate day-0 procurement) rather than an allocation feature like economic efficiency, and the allowed mechanisms are restricted to those judged to require only an incremental modification of the current game. Replacing the supply-chain negotiation procedures with a one-shot direct mechanism, for example, was not an option. We believe that such operational restrictions and idiosyncratic objectives are actually quite typical of practical mechanism design settings, where they are perhaps more commonly characterized as incentive engineering problems.

In response to the problem, the TAC/SCM designers adopted several rule changes intended to penalize large day-0 orders. These included modifications to supplier pricing policies and introduction of storage costs assessed on inventories of components and finished goods. Despite the changes, day-0 procurement was very high in the early rounds of the 2004 competition. In a drastic measure, the GameMaster imposed a fivefold increase of storage costs midway through the tournament. Even this did not stem the tide, and day-0 procurement in the final rounds actually *increased* (by some measures) from 2003 [9].

The apparent difficulty in identifying rule modifications that effect moderation in day-0 procurement is quite striking. Although the designs were widely discussed, predictions for the effects of various proposals were supported primarily by intuitive arguments or at best by back-of-the-envelope calculations. Much of the difficulty, of course, is anticipating the agents' (and their developers') responses without essentially running a gaming exercise for this purpose. The episode caused us to consider whether new ap-

[1]Information about TAC and the SCM game, including specifications, rules, and competition results, can be found at `http://www.sics.se/tac`.

EC'06, June 11–15, 2006, Ann Arbor, Michigan, USA.

proaches or tools could enable more systematic analysis of design options. Standard game-theoretic and mechanism design methods are clearly relevant, although the lack of an analytic description of the game seems to be an impediment. Under the assumption that the simulator itself is the only reliable source of outcome computation, we refer to our task as *empirical* mechanism design.

In the sequel, we develop some general methods for empirical mechanism design and apply them to the TAC/SCM redesign problem. Our analysis focuses on the setting of storage costs (taking other game modifications as fixed), since this is the most direct deterrent to early procurement adopted. Our results confirm the basic intuition that incentives for day-0 purchasing decrease as storage costs rise. We also confirm that the high day-0 procurement observed in the 2004 tournament is a rational response to the setting of storage costs used. Finally, we conclude from our data that it is very unlikely that any reasonable setting of storage costs would result in acceptable levels of day-0 procurement, so a different design approach would have been required to eliminate this problem.

Overall, we contribute a formal framework and a set of methods for tackling indirect mechanism design problems in settings where only a black-box description of players' utilities is available. Our methods incorporate estimation of *sets* of Nash equilibria and sample Nash equilibria, used in conjuction to support general claims about the structure of the mechanism designer's utility, as well as a restricted probabilistic analysis to assess the likelihood of conclusions. We believe that most realistic problems are too complex to be amenable to exact analysis. Consequently, we advocate the approach of gathering evidence to provide indirect support of specific hypotheses.

2. PRELIMINARIES

A *normal form game*[2] is denoted by $[I, \{R_i\}, \{u_i(r)\}]$, where I refers to the set of players and $m = |I|$ is the number of players. R_i is the set of strategies available to player $i \in I$, with $R = R_1 \times \ldots \times R_m$ representing the set of joint strategies of all players. We designate the set of pure strategies available to player i by A_i, and denote the joint set of pure strategies of all players by $A = A_1 \times \ldots \times A_m$. It is often convenient to refer to a strategy of player i separately from that of the remaining players. To accommodate this, we use a_{-i} to denote the joint strategy of all players other than player i.

Let S_i be the set of all probability distributions (mixtures) over A_i and, similarly, S be the set of all distributions over A. An $s \in S$ is called a mixed strategy profile. When the game is finite (i.e., A and I are both finite), the probability that $a \in A$ is played under s is written $s(a) = s(a_i, a_{-i})$. When the distribution s is not correlated, we can simply say $s_i(a_i)$ when referring to the probability player i plays a_i under s.

Next, we define the payoff (utility) function of each player i by $u_i : A_1 \times \cdots \times A_m \to \mathbb{R}$, where $u_i(a_i, a_{-i})$ indicates the payoff to player i to playing pure strategy a_i when the remaining players play a_{-i}. We can extend this definition to mixed strategies by assuming that u_i are von Neumann-Morgenstern (vNM) utilities as follows: $u_i(s) = E_s[u_i]$, where E_s is the expectation taken with respect to the probability distribution of play induced by the players' mixed strategy s. When the game has finitely many pure strategies, this definition is equivalent to $u_i(s) = \sum_{a \in A} u_i(a)s(a)$.

Occasionally, we write $u_i(x, y)$ to mean that $x \in A_i$ or S_i and $y \in A_{-i}$ or S_{-i} depending on context. We also express the set of utility functions of all players as $u(\cdot) = \{u_1(\cdot), \ldots, u_m(\cdot)\}$.

We define a function, $\epsilon : R \to \mathbb{R}$, interpreted as the maximum benefit any player can obtain by deviating from its strategy in the specified profile.

$$\epsilon(r) = \max_{i \in I} \max_{a_i \in A_i} [u_i(a_i, r_{-i}) - u_i(r)], \tag{1}$$

where r belongs to some strategy set, R, of either pure or mixed strategies.

Faced with a game, an agent would ideally play its best strategy given those played by the other agents. A configuration where all agents play strategies that are best responses to the others constitutes a *Nash equilibrium*.

DEFINITION 1. *A strategy profile* $r = (r_1, \ldots, r_m)$ *constitutes a* Nash equilibrium *of game* $[I, \{R_i\}, \{u_i(r)\}]$ *if for every* $i \in I$, $r_i' \in R_i$, $u_i(r_i, r_{-i}) \geq u_i(r_i', r_{-i})$.

When $r \in A$, the above defines a *pure strategy Nash equilibrium*; otherwise the definition describes a *mixed strategy Nash equilibrium*. We often appeal to the concept of an *approximate*, or *ϵ-Nash equilibrium*, where ϵ is the maximum benefit to any agent for deviating from the prescribed strategy. Thus, $\epsilon(r)$ as defined above (1) is such that profile r is an ϵ-Nash equilibrium iff $\epsilon(r) \leq \epsilon$.

In this study we devote particular attention to games that exhibit symmetry with respect to payoffs, rendering agents strategically identical.

DEFINITION 2. *A game* $[I, \{R_i\}, \{u_i(r)\}]$ *is* symmetric *if for all* $i, j \in I$, *(a)* $R_i = R_j$ *and (b)* $u_i(r_i, r_{-i}) = u_j(r_j, r_{-j})$ *whenever* $r_i = r_j$ *and* $r_{-i} = r_{-j}$

3. THE MODEL

We model the strategic interactions between the designer of the mechanism and its participants as a two-stage game. The designer moves first by selecting a value, θ, from a set of allowable mechanism settings, Θ. All the participant agents observe the mechanism parameter θ and move simultaneously thereafter. For example, the designer could be deciding between a first-price and second-price sealed-bid auction mechanisms, with the presumption that after the choice has been made, the bidders will participate with full awareness of the auction rules.

Since the participants play with full knowledge of the mechanism parameter, we define a game between them in the second stage as $\Gamma_\theta = [I, \{R_i\}, \{u_i(r, \theta)\}]$. We refer to Γ_θ as a game *induced* by θ. Let $\mathcal{N}(\theta)$ be the set of strategy profiles considered *solutions* of the game Γ_θ.[3]

Suppose that the goal of the designer is to optimize the value of some welfare function, $W(r, \theta)$, dependent on the mechanism parameter and resulting play, r. We define a pessimistic measure, $W(\hat{R}, \theta) = \inf\{W(r, \theta) : r \in \hat{R}\}$, representing the worst-case welfare of the game induced by θ, assuming that agents play some joint strategy in $\hat{R}$. Typically we care about $W(\mathcal{N}(\theta), \theta)$, the worst-case outcome of playing *some* solution.[4]

[2]By employing the normal form, we model agents as playing a single action, with decisions taken simultaneously. This is appropriate for our current study, which treats strategies (agent programs) as atomic actions. We could capture finer-grained decisions about action over time in the *extensive form*. Although any extensive game can be recast in normal form, doing so may sacrifice compactness and blur relevant distinctions (e.g., subgame perfection).

[3]We generally adopt Nash equilibrium as the solution concept, and thus take $\mathcal{N}(\theta)$ to be the set of equilibria. However, much of the methodology developed here could be employed with alternative criteria for deriving agent behavior from a game definition.

[4]Again, alternatives are available. For example, if one has a probability distribution over the solution set $\mathcal{N}(\theta)$, it would be natural to take the expectation of $W(r, \theta)$ instead.

On some problems we can gain considerable advantage by using an *aggregation function* to map the welfare outcome of a game specified in terms of agent strategies to an equivalent welfare outcome specified in terms of a lower-dimensional summary.

DEFINITION 3. *A function* $\phi : R_1 \times \cdots \times R_m \to \mathbb{R}^q$ *is an* aggregation function *if* $m \geq q$ *and* $W(r, \theta) = V(\phi(r), \theta)$ *for some function* V.

We overload the function symbol to apply to *sets* of strategy profiles: $\phi(\hat{R}) = \{\phi(r) : r \in \hat{R}\}$. For convenience of exposition, we write $\phi^*(\theta)$ to mean $\phi(\mathcal{N}(\theta))$.

Using an aggregation function yields a more compact representation of strategy profiles. For example, suppose—as in our application below—that an agent's strategy is defined by a numeric parameter. If all we care about is the total value played, we may take $\phi(a) = \sum_{i=1}^{m} a_i$. If we have chosen our aggregator carefully, we may also capture structure not obvious otherwise. For example, $\phi^*(\theta)$ could be decreasing in θ, whereas $\mathcal{N}(\theta)$ might have a more complex structure.

Given a description of the solution correspondence $\mathcal{N}(\theta)$ (equivalently, $\phi^*(\theta)$), the designer faces a standard optimization problem. Alternatively, given a simulator that could produce an unbiased sample from the distribution of $W(\mathcal{N}(\theta), \theta)$ for any θ, the designer would be faced with another much appreciated problem in the literature: simulation optimization [12].

However, even for a game Γ_θ with known payoffs it may be computationally intractable to solve for Nash equilibria, particularly if the game has large or infinite strategy sets. Additionally, we wish to study games where the payoffs are not explicitly given, but must be determined from simulation or other experience with the game.[5]

Accordingly, we assume that we are given a (possibly noisy) data set of payoff realizations: $\mathcal{D}_o = \{(\theta^1, a^1, U^1), \ldots, (\theta^k, a^k, U^k)\}$, where for every data point θ^i is the observed mechanism parameter setting, a^i is the observed pure strategy profile of the participants, and U^i is the corresponding realization of agent payoffs. We may also have additional data generated by a (possibly noisy) simulator: $\mathcal{D}_s = \{(\theta^{k+1}, a^{k+1}, U^{k+1}), \ldots, (\theta^{k+l}, a^{k+l}, U^{k+l})\}$. Let $\mathcal{D} = \{\mathcal{D}_o, \mathcal{D}_s\}$ be the combined data set. (Either $\mathcal{D}_o$ or $\mathcal{D}_s$ may be null for a particular problem.)

In the remainder of this paper, we apply our modeling approach, together with several empirical game-theoretic methods, in order to answer questions regarding the design of the TAC/SCM scenario.

4. EMPIRICAL DESIGN ANALYSIS

Since our data comes in the form of payoff experience and not as the value of an objective function for given settings of the control variable, we can no longer rely on the methods for optimizing functions using simulations. Indeed, a fundamental aspect of our design problem involves estimating the Nash equilibrium correspondence. Furthermore, we cannot rely directly on the convergence results that abound in the simulation optimization literature, and must establish probabilistic analysis methods tailored for our problem setting.

4.1 TAC/SCM Design Problem

We describe our empirical design analysis methods by presenting a detailed application to the TAC/SCM scenario introduced above. Recall that during the 2004 tournament, the designers of the supply-chain game chose to dramatically increase storage costs as a measure aimed at curbing day-0 procurement, to little avail. Here we systematically explore the relationship between storage costs and the aggregate quantity of components procured on day 0 in equilibrium. In doing so, we consider several questions raised during and after the tournament. First, does increasing storage costs actually reduce day-0 procurement? Second, was the excessive day-0 procurement that was observed during the 2004 tournament rational? And third, could increasing storage costs sufficiently have reduced day-0 procurement to an "acceptable" level, and if so, what should the setting of storage costs have been? It is this third question that defines the mechanism design aspect of our analysis.[6]

To apply our methods, we must specify the agent strategy sets, the designer's welfare function, the mechanism parameter space, and the source of data. We restrict the agent strategies to be a multiplier on the quantity of the day-0 requests by one of the finalists, Deep Maize, in the 2004 TAC/SCM tournament. We further restrict it to the set [0,1.5], since any strategy below 0 is illegal and strategies above 1.5 are extremely aggressive (thus unlikely to provide refuting deviations beyond those available from included strategies, and certainly not part of any desirable equilibrium). All other behavior is based on the behavior of Deep Maize and is identical for all agents. This choice can provide only an estimate of the actual tournament behavior of a "typical" agent. However, we believe that the general form of the results should be robust to changes in the full agent behavior.

We model the designer's welfare function as a threshold on the sum of day-0 purchases. Let $\phi(a) = \sum_{i=1}^{6} a_i$ be the aggregation function representing the sum of day-0 procurement of the six agents participating in a particular supply-chain game (for mixed strategy profiles s, we take expectation of ϕ with respect to the mixture). The designer's welfare function $W(\mathcal{N}(\theta), \theta)$ is then given by $\mathbf{I}\{\sup\{\phi^*(\theta)\} \leq \alpha\}$, where α is the maximum acceptable level of day-0 procurement and $\mathbf{I}$ is the indicator function. The designer selects a value θ of storage costs, expressed as an annual percentage of the baseline value of components in the inventory (charged daily), from the set $\Theta = \mathbb{R}^+$. Since the designer's decision depends only on $\phi^*(\theta)$, we present all of our results in terms of the value of the aggregation function.

4.2 Estimating Nash Equilibria

The objective of TAC/SCM agents is to maximize profits realized over a game instance. Thus, if we fix a strategy for each agent at the beginning of the simulation and record the corresponding profits at the end, we will have obtained a data point in the form $(a, U(a))$. If we also have fixed the parameter θ of the simulator, the resulting data point becomes part of our data set $\mathcal{D}$. This data set, then, contains data only in the form of pure strategies of players and their corresponding payoffs, and, consequently, in order to formulate the designer's problem as optimization, we must first determine or approximate the set of Nash equilibria of each game Γ_θ. Thus, we need methods for approximating Nash equilibria for infinite games. Below, we describe the two methods we used in our study. The first has been explored empirically before, whereas the second is introduced here as the method specifically designed to approximate *a set* of Nash equilibria.

4.2.1 *Payoff Function Approximation*

The first method for estimating Nash equilibria based on data

[5]This is often the case for real games of interest, where natural language or algorithmic descriptions may substitute for a formal specification of strategy and payoff functions.

[6]We do not address whether and how other measures (e.g., constraining procurement directly) could have achieved design objectives. Our approach takes as given some set of design options, in this case defined by the storage cost parameter. In principle our methods could be applied to a different or larger design space, though with corresponding complexity growth.

uses supervised learning to approximate payoff functions of mechanism participants from a data set of game experience [17]. Once approximate payoff functions are available for all players, the Nash equilibria may be either found analytically or approximated using numerical techniques, depending on the learning model. In what follows, we estimate only a sample Nash equilibrium using this technique, although this restriction can be removed at the expense of additional computation time.

One advantage of this method is that it can be applied to any data set and does not require the use of a simulator. Thus, we can apply it when $\mathcal{D}_s = \emptyset$. If a simulator is available, we can generate additional data to build confidence in our initial estimates.[7]

We tried the following methods for approximating payoff functions: quadratic regression (QR), locally weighted average (LWA), and locally weighted linear regression (LWLR). We also used control variates to reduce the variance of payoff estimates, as in our previous empirical game-theoretic analysis of TAC/SCM-03 [19].

The quadratic regression model makes it possible to compute equilibria of the learned game analytically. For the other methods we applied replicator dynamics [7] to a discrete approximation of the learned game. The expected total day-0 procurement in equilibrium was taken as the estimate of an outcome.

4.2.2 Search in Strategy Profile Space

When we have access to a simulator, we can also use directed search through profile space to estimate the set of Nash equilibria, which we describe here after presenting some additional notation.

DEFINITION 4. *A* strategic neighbor *of a pure strategy profile a is a profile that is identical to a in all but one strategy. We define $S_{nb}(a, \mathcal{D})$ as the set of all strategic neighbors of a available in the data set $\mathcal{D}$. Similarly, we define $S_{nb}(a, \tilde{\mathcal{D}})$ to be all strategic neighbors of a not in $\mathcal{D}$. Finally, for any $a' \in S_{nb}(a, \mathcal{D})$ we define the deviating agent as $i(a, a')$.*

DEFINITION 5. *The ϵ-bound, $\hat{\epsilon}$, of a pure strategy profile a is defined as $\max_{a' \in S_{nb}(a,\mathcal{D})} \max\{u_{i(a,a')}(a') - u_{i(a,a')}(a), 0\}$. We say that a is a* candidate δ-equilibrium *for $\delta \geq \hat{\epsilon}$.*

When $S_{nb}(a, \tilde{\mathcal{D}}) = \emptyset$ (i.e., all strategic neighbors are represented in the data), a is *confirmed* as an $\hat{\epsilon}$-Nash equilibrium.

Our search method operates by exploring deviations from candidate equilibria. We refer to it as "BestFirstSearch", as it selects with probability one a strategy profile $a' \in S_{nb}(a, \tilde{\mathcal{D}})$ that has the smallest $\hat{\epsilon}$ in $\mathcal{D}$.

Finally we define an estimator for a set of Nash equilibria.

DEFINITION 6. *For a set K, define $Co(K)$ to be the convex hull of K. Let B_δ be the set of candidates at level δ. We define $\hat{\phi}^*(\theta) = Co(\{\phi(a) : a \in B_\delta\})$ for a fixed δ to be an estimator of $\phi^*(\theta)$.*

In words, the estimate of a set of equilibrium outcomes is the convex hull of all aggregated strategy profiles with ϵ-bound below some fixed δ. This definition allows us to exploit structure arising from the aggregation function. If two profiles are close in terms of aggregation values, they may be likely to have similar ϵ-bounds. In particular, if one is an equilibrium, the other may be as well. We present some theoretical support for this method of estimating the set of Nash equilibria below.

Since the game we are interested in is infinite, it is necessary to terminate BestFirstSearch before exploring the entire space of strategy profiles. We currently determine termination time in a somewhat ad hoc manner, based on observations about the current set of candidate equilibria.[8]

4.3 Data Generation

Our data was collected by simulating TAC/SCM games on a local version of the 2004 TAC/SCM server, which has a configuration setting for the storage cost. Agent strategies in simulated games were selected from the set $\{0, 0.3, 0.6, \ldots, 1.5\}$ in order to have positive probability of generating strategic neighbors.[9] A baseline data set $\mathcal{D}_o$ was generated by sampling 10 randomly generated strategy profiles for each $\theta \in \{0, 50, 100, 150, 200\}$. Between 5 and 10 games were run for each profile after discarding games that had various flaws.[10] We used search to generate a simulated data set $\mathcal{D}_s$, performing between 12 and 32 iterations of BestFirstSearch for each of the above settings of θ. Since simulation cost is extremely high (a game takes nearly 1 hour to run), we were able to run a total of 2670 games over the span of more than six months. For comparison, to get the entire description of an empirical game defined by the restricted finite joint strategy space for each value of $\theta \in \{0, 50, 100, 150, 200\}$ would have required at least 23100 games (sampling each profile 10 times).

4.4 Results

4.4.1 Analysis of the Baseline Data Set

We applied the three learning methods described above to the baseline data set $\mathcal{D}_o$. Additionally, we generated an estimate of the Nash equilibrium correspondence, $\hat{\phi}^*(\theta)$, by applying Definition 6 with δ =2.5E6. The results are shown in Figure 1. As we can see, the correspondence $\hat{\phi}^*(\theta)$ has little predictive power based on $\mathcal{D}_o$, and reveals no interesting structure about the game. In contrast, all three learning methods suggest that total day-0 procurement is a decreasing function of storage costs.

4.4.2 Analysis of Search Data

To corroborate the initial evidence from the learning methods, we estimated $\hat{\phi}^*(\theta)$ (again, using δ =2.5E6) on the data set $D = \{\mathcal{D}_o, \mathcal{D}_s\}$, where $\mathcal{D}_s$ is data generated through the application of BestFirstSearch. The results of this estimate are plotted against the results of the learning methods trained on $\mathcal{D}_o$[11] in Figure 2. First, we note that the addition of the search data narrows the range of potential equilibria substantially. Furthermore, the actual point predictions of the learning methods and those based on ϵ-bounds after search are reasonably close. Combining the evidence gathered from these two very different approaches to estimating the outcome correspondence yields a much more compelling picture of the relationship between storage costs and day-0 procurement than either method used in isolation.

[7]For example, we can use active learning techniques [5] to improve the quality of payoff function approximation. In this work, we instead concentrate on search in strategy profile space.

[8]Generally, search is terminated once the set of candidate equilibria is small enough to draw useful conclusions about the likely range of equilibrium strategies in the game.

[9]Of course, we do not restrict our Nash equilibrium estimates to stay in this discrete subset of [0,1.5].

[10]For example, if we detected that any agent failed during the game (failures included crashes, network connectivity problems, and other obvious anomalies), the game would be thrown out.

[11]It is unclear how meaningful the results of learning would be if D_s were added to the training data set. Indeed, the additional data may actually increase the learning variance.

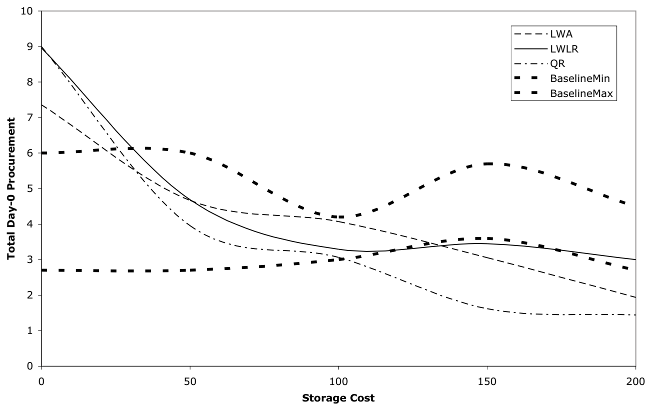

Figure 1: Aggregate day-0 procurement estimates based on $\mathcal{D}_o$. The correspondence $\hat{\phi}^*(\theta)$ is the interval between "BaselineMin" and "BaselineMax".

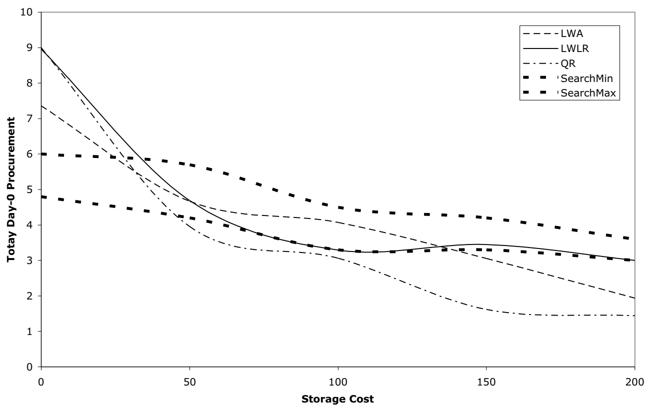

Figure 2: Aggregate day-0 procurement estimates based on search in strategy profile space compared to function approximation techniques trained on $\mathcal{D}_o$. The correspondence $\hat{\phi}^*(\theta)$ for $\mathcal{D} = \{\mathcal{D}_o, \mathcal{D}_s\}$ is the interval between "SearchMin" and "SearchMax".

This evidence supports the initial intuition that day-0 procurement should be decreasing with storage costs. It also confirms that high levels of day-0 procurement are a rational response to the 2004 tournament setting of average storage cost, which corresponds to $\theta = 100$. The *minimum* prediction for aggregate procurement at this level of storage costs given by any experimental methods is approximately 3. This is quite high, as it corresponds to an expected commitment of 1/3 of the total supplier capacity for the entire game. The maximum prediction is considerably higher at 4.5. In the actual 2004 competition, aggregate day-0 procurement was equivalent to 5.71 on the scale used here [9]. Our predictions underestimate this outcome to some degree, but show that any rational outcome was likely to have high day-0 procurement.

4.4.3 Extrapolating the Solution Correspondence

We have reasonably strong evidence that the outcome correspondence is decreasing. However, the ultimate goal is to be able to either set the storage cost parameter to a value that would curb day-0 procurement in equilibrium or conclude that this is not possible.

To answer this question directly, suppose that we set a conservative threshold $\alpha = 2$ on aggregate day-0 procurement.[12] Linear extrapolation of the maximum of the outcome correspondence estimated from $\mathcal{D}$ yields $\theta = 320$.

The data for $\theta = 320$ were collected in the same way as for other storage cost settings, with 10 randomly generated profiles followed by 33 iterations of BestFirstSearch. Figure 3 shows the detailed ϵ-bounds for all profiles in terms of their corresponding values of ϕ.

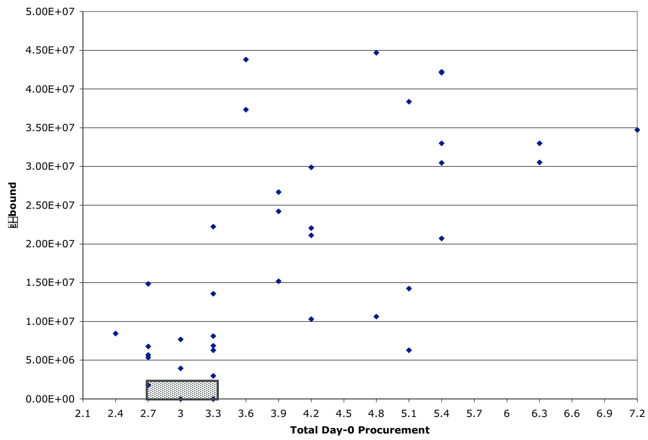

Figure 3: Values of $\hat{\epsilon}$ for profiles explored using search when $\theta = 320$. Strategy profiles explored are presented in terms of the corresponding values of $\phi(a)$. The gray region corresponds to $\hat{\phi}^*(320)$ with δ =2.5M.

The estimated set of aggregate day-0 outcomes is very close to that for $\theta = 200$, indicating that there is little additional benefit to raising storage costs above 200. Observe, that even the lower bound of our estimated set of Nash equilibria is well above the target day-0 procurement of 2. Furthermore, payoffs to agents are almost always negative at $\theta = 320$. Consequently, increasing the costs further would be undesirable even if day-0 procurement could eventually be curbed. Since we are reasonably confident that $\phi^*(\theta)$ is decreasing in θ, we also do not expect that setting θ somewhere between 200 and 320 will achieve the desired result.

We conclude that it is unlikely that day-0 procurement could ever be reduced to a desirable level using any reasonable setting of the storage cost parameter. That our predictions tend to underestimate tournament outcomes reinforces this conclusion. To achieve the desired reduction in day-0 procurement requires redesigning other aspects of the mechanism.

4.5 Probabilistic Analysis

Our empirical analysis has produced evidence in support of the conclusion that no reasonable setting of storage cost was likely to sufficiently curb excessive day-0 procurement in TAC/SCM '04. All of this evidence has been in the form of simple interpolation and extrapolation of estimates of the Nash equilibrium correspondence. These estimates are based on simulating game instances, and are subject to sampling noise contributed by the various stochastic elements of the game. In this section, we develop and apply methods for evaluating the sensitivity of our ϵ-bound calculations to such stochastic effects.

[12]Recall that designer's objective is to incentivize aggergate day-0 procurement that is below the threshold α. Our threshold here still represents a commitment of over 20% of the suppliers' capacity for the entire game on average, so in practice we would probably want the threshold to be even lower.

Suppose that all agents have finite (and small) pure strategy sets, A. Thus, it is feasible to sample the entire payoff matrix of the game. Additionally, suppose that noise is additive with zero-mean and finite variance, that is, $U_i(a) = u_i(a) + \tilde{\xi}_i(a)$, where $U_i(a)$ is the observed payoff to i when a was played, $u_i(a)$ is the actual corresponding payoff, and $\tilde{\xi}_i(a)$ is a mean-zero normal random variable. We designate the known variance of $\tilde{\xi}_i(a)$ by $\sigma_i^2(a)$. Thus, we assume that $\tilde{\xi}_i(a)$ is normal with distribution $N(0, \sigma_i^2(a))$.

We take $\bar{u}_i(a)$ to be the sample mean over all $U_i(a)$ in $\mathcal{D}$, and follow Chang and Huang [3] to assume that we have an improper prior over the actual payoffs $u_i(a)$ and sampling was independent for all i and a. We also rely on their result that $u_i(a)|\bar{u}_i(a) = \bar{u}_i(a) - Z_i(a)/[\sigma_i(a)/\sqrt{n_i(a)}]$ are independent with posterior distributions $N(\bar{u}_i(a), \sigma_i^2(a)/n_i(a))$, where $n_i(a)$ is the number of samples taken of payoffs to i for pure profile a, and $Z_i(a) \sim N(0,1)$.

We now derive a generic probabilistic bound that a profile $a \in A$ is an ϵ-Nash equilibrium. If $u_i(\cdot)|\bar{u}_i(\cdot)$ are independent for all $i \in I$ and $a \in A$, we have the following result (from this point on we omit conditioning on $\bar{u}_i(\cdot)$ for brevity):

PROPOSITION 1.

$$\Pr\left(\max_{i\in I}\max_{b\in A_i} u_i(b, a_{-i}) - u_i(a) \le \epsilon\right) = \\ = \prod_{i\in I}\int_{\mathbb{R}} \prod_{b\in A_i\setminus a_i} \Pr(u_i(b, a_{-i}) \le u + \epsilon) f_{u_i(a)}(u)du, \tag{2}$$

where $f_{u_i(a)}(u)$ is the pdf of $N(\bar{u}_i(a), \sigma_i(a))$.

The proofs of this and all subsequent results are in the Appendix.

The posterior distribution of the optimum mean of n samples, derived by Chang and Huang [3], is

$$\Pr(u_i(a) \le c) = 1 - \Phi\left[\frac{\sqrt{n_i(a)}(\bar{u}_i(a) - c)}{\sigma_i(a)}\right], \tag{3}$$

where $a \in A$ and $\Phi(\cdot)$ is the $N(0,1)$ distribution function.

Combining the results (2) and (3), we obtain a probabilistic confidence bound that $\epsilon(a) \le \gamma$ for a given γ.

Now, we consider cases of incomplete data and use the results we have just obtained to construct an upper bound (restricted to profiles represented in data) on the distribution of $\sup\{\phi^*(\theta)\}$ and $\inf\{\phi^*(\theta)\}$ (assuming that both are attainable):

$$\Pr\{\sup\{\phi^*(\theta)\} \le x\} \le_{\mathcal{D}} \\ \Pr\{\exists a \in \mathcal{D} : \phi(a) \le x \wedge a \in \mathcal{N}(\theta)\} \le \\ \sum_{a\in\mathcal{D}:\phi(a)\le x} \Pr\{a \in \mathcal{N}(\theta)\} = \sum_{a\in\mathcal{D}:\phi(a)\le x} \Pr\{\epsilon(a) = 0\},$$

where x is a real number and $\le_{\mathcal{D}}$ indicates that the upper bound accounts only for strategies that appear in the data set $\mathcal{D}$. Since the events $\{\exists a \in \mathcal{D} : \phi(a) \le x \wedge a \in \mathcal{N}(\theta)\}$ and $\{\inf\{\phi^*(\theta)\} \le x\}$ are equivalent, this also defines an upper bound on the probability of $\{\inf\{\phi^*(\theta)\} \le x\}$. The values thus derived comprise the Tables 1 and 2.

Tables 1 and 2 suggest that the existence of *any* equilibrium with $\phi(a) < 2.7$ is unlikely for any θ that we have data for, although this judgment, as we mentioned, is only with respect to the profiles we have actually sampled. We can then accept this as another piece of evidence that the designer could not find a suitable setting of θ to achieve his objectives—indeed, the designer seems unlikely to achieve his objective even if he could persuade participants to play a desirable equilibrium!

$\phi^*(\theta)$	$\theta = 0$	$\theta = 50$	$\theta = 100$
<2.7	0.000098	0	0.146
<3	0.158	0.0511	0.146
<3.9	0.536	0.163	1
<4.5	1	1	1

Table 1: Upper bounds on the distribution of $\inf\{\phi^*(\theta)\}$ restricted to $\mathcal{D}$ for $\theta \in \{0, 50, 100\}$ when $\mathcal{N}(\theta)$ is a set of Nash equilibria.

$\phi^*(\theta)$	$\theta = 150$	$\theta = 200$	$\theta = 320$
<2.7	0	0	0.00132
<3	0.0363	0.141	1
<3.9	1	1	1
<4.5	1	1	1

Table 2: Upper bounds on the distribution of $\inf\{\phi^*(\theta)\}$ restricted to $\mathcal{D}$ for $\theta \in \{150, 200, 320\}$ when $\mathcal{N}(\theta)$ is a set of Nash equilibria.

Table 1 also provides additional evidence that the agents in the 2004 TAC/SCM tournament were indeed rational in procuring large numbers of components at the beginning fo the game. If we look at the third column of this table, which corresponds to $\theta = 100$, we can gather that no profile a in our data with $\phi(a) < 3$ is very likely to be played in equilibrium.

The bounds above provide some general evidence, but ultimately we are interested in a concrete probabilistic assessment of our conclusion with respect to the data we have sampled. Particularly, we would like to say something about what happens for the settings of θ for which we have no data. To derive an approximate probabilistic bound on the probability that no $\theta \in \Theta$ could have achieved the designer's objective, let $\cup_{j=1}^{J}\Theta_j$, be a partition of Θ, and assume that the function $\sup\{\phi^*(\theta)\}$ satisfies the *Lipschitz condition* with *Lipschitz constant* A_j on each subset Θ_j.[13] Since we have determined that raising the storage cost above 320 is undesirable due to secondary considerations, we restrict attention to $\Theta = [0, 320]$. We now define each subset j to be the interval between two points for which we have produced data. Thus,

$$\Theta = [0, 50]\bigcup(50, 100]\bigcup(100, 150]\bigcup(150, 200]\bigcup(200, 320],$$

with j running between 1 and 5, corresponding to subintervals above. We will further denote each Θ_j by $(a_j, b_j]$.[14] Then, the following Proposition gives us an *approximate* upper bound[15] on the probability that $\sup\{\phi^*(\theta)\} \le \alpha$.

PROPOSITION 2.

$$\Pr\{\bigvee_{\theta\in\Theta} \sup\{\phi(\theta)\} \le \alpha\} \le_{\mathcal{D}} \\ \sum_{j=1}^{5} \sum_{y,z\in D:y+z\le c_j} \left(\sum_{a:\phi(a)=z} \Pr\{\epsilon(a) = 0\}\right) \times \\ \times \left(\sum_{a:\phi(a)=y} \Pr\{\epsilon(a) = 0\}\right),$$

[13] A function that satisfies the *Lipschitz condition* is called *Lipschitz continuous*.

[14] The treatment for the interval [0,50] is identical.

[15] It is approximate in a sense that we only take into account strategies that are present in the data.

where $c_j = 2\alpha + A_j(b_j - a_j)$ and $\leq_{\mathcal{D}}$ indicates that the upper bound only accounts for strategies that appear in the data set $\mathcal{D}$.

Due to the fact that our bounds are approximate, we cannot use them as a conclusive probabilistic assessment. Instead, we take this as another piece of evidence to complement our findings.

Even if we can assume that a function that we approximate from data is Lipschitz continuous, we rarely actually know the Lipschitz constant for any subset of Θ. Thus, we are faced with a task of estimating it from data. Here, we tried three methods of doing this. The first one simply takes the highest slope that the function attains within the available data and uses this constant value for every subinterval. This produces the most conservative bound, and in many situations it is unlikely to be informative.

An alternative method is to take an upper bound on slope obtained *within each subinterval* using the available data. This produces a much less conservative upper bound on probabilities. However, since the actual upper bound is generally greater for each subinterval, the resulting probabilistic bound may be deceiving.

A final method that we tried is a compromise between the two above. Instead of taking the conservative upper bound based on data over the entire function domain Θ, we take the average of upper bounds obtained at each Θ_j. The bound at an interval is then taken to be the maximum of the upper bound for this interval and the average upper bound for all intervals.

The results of evaluating the expression for

$$\Pr\{\bigvee_{\theta\in\Theta} \sup\{\phi^*(\theta)\} \leq \alpha\}$$

when $\alpha = 2$ are presented in Table 3. In terms of our claims in

$\max_j A_j$	A_j	$\max\{A_j,\text{ave}(A_j)\}$
1	0.00772	0.00791

Table 3: Approximate upper bound on probability that some setting of $\theta \in [0, 320]$ will satisfy the designer objective with target $\alpha = 2$. Different methods of approximating the upper bound on slope in each subinterval j are used.

this work, the expression gives an upper bound on the probability that some setting of θ (i.e., storage cost) in the interval [0,320] will result in total day-0 procurement that is no greater in any equilibrium than the target specified by α and taken here to be 2. As we had suspected, the most conservative approach to estimating the upper bound on slope, presented in the first column of the table, provides us little information here. However, the other two estimation approaches, found in columns two and three of Table 3, suggest that we are indeed quite confident that no reasonable setting of $\theta \in [0, 320]$ would have done the job. Given the tremendous difficulty of the problem, this result is very strong.[16] Still, we must be very cautious in drawing too heroic a conclusion based on this evidence. Certainly, we have not "checked" all the profiles but only a small proportion of them (infinitesimal, if we consider the entire continuous domain of θ and strategy sets). Nor can we expect ever to obtain enough evidence to make completely objective conclusions. Instead, the approach we advocate here is to collect as much evidence as is feasible given resource constraints, and make the most compelling judgment based on this evidence, if at all possible.

[16]Since we did not have all the possible deviations for any profile available in the data, the true upper bounds may be even lower.

5. CONVERGENCE RESULTS

At this point, we explore abstractly whether a design parameter choice based on payoff data can be asymptotically reliable.

As a matter of convenience, we will use notation $u_{n,i}(a)$ to refer to a payoff function of player i based on an average over n i.i.d. samples from the distribution of payoffs. We also assume that $u_{n,i}(a)$ are independent for all $a \in A$ and $i \in I$. We will use the notation Γ_n to refer to the game $[I, R, \{u_{i,n}(\cdot)\}]$, whereas Γ will denote the "underlying" game, $[I, R, \{u_i(\cdot)\}]$. Similarly, we define $\epsilon_n(r)$ to be $\epsilon(r)$ with respect to the game Γ_n.

In this section, we show that $\epsilon_n(s) \to \epsilon(s)$ a.s. uniformly on the mixed strategy space for any finite game, and, furthermore, that all mixed strategy Nash equilibria in empirical games eventually become arbitrarily close to *some* Nash equilibrium strategies in the underlying game. We use these results to show that under certain conditions, the optimal choice of the design parameter based on empirical data converges almost surely to the actual optimum.

THEOREM 3. *Suppose that $|I| < \infty, |A| < \infty$. Then $\epsilon_n(s) \to \epsilon(s)$ a.s. uniformly on S.*

Recall that $\mathcal{N}$ is a set of all Nash equilibria of Γ. If we define $\mathcal{N}_{n,\gamma} = \{s \in S : \epsilon_n(s) \leq \gamma\}$, we have the following corollary to Theorem 3:

COROLLARY 4. *For every $\gamma > 0$, there is M such that $\forall n \geq M$, $\mathcal{N} \subset \mathcal{N}_{n,\gamma}$ a.s.*

PROOF. Since $\epsilon(s) = 0$ for every $s \in \mathcal{N}$, we can find M large enough such that $\Pr\{\sup_{n\geq M} \sup_{s\in\mathcal{N}} \epsilon_n(s) < \gamma\} = 1$. □

By the Corollary, for any game with a finite set of pure strategies and for any $\epsilon > 0$, all Nash equilibria lie in the set of empirical ϵ-Nash equilibria if enough samples have been taken. As we now show, this provides some justification for our use of a set of profiles with a non-zero ϵ-bound as an estimate of the set of Nash equilibria.

First, suppose we conclude that for a particular setting of θ, $\sup\{\hat{\phi}^*(\theta)\} \leq \alpha$. Then, since for any fixed $\epsilon > 0$, $\mathcal{N}(\theta) \subset \mathcal{N}_{n,\epsilon}(\theta)$ when n is large enough,

$$\sup\{\phi^*(\theta)\} = \sup_{s\in\mathcal{N}(\theta)} \phi(s) \leq \sup_{s\in\mathcal{N}_{n,\epsilon}(\theta)} \phi(s) = \sup\{\hat{\phi}^*(\theta)\} \leq \alpha$$

for any such n. Thus, since we defined the welfare function of the designer to be $\mathbf{I}\{\sup\{\phi^*(\theta)\} \leq \alpha\}$ in our domain of interest, the empirical choice of θ satisfies the designer's objective, thereby maximizing his welfare function.

Alternatively, suppose we conclude that $\inf\{\hat{\phi}^*(\theta)\} > \alpha$ for every θ in the domain. Then,

$$\alpha < \inf\{\hat{\phi}^*(\theta)\} = \inf_{s\in\mathcal{N}_{n,\epsilon}(\theta)} \phi(s) \leq \inf_{s\in\mathcal{N}(\theta)} \phi(s) \leq \leq \sup_{s\in\mathcal{N}(\theta)} \phi(s) = \sup\{\phi^*(\theta)\},$$

for every θ, and we can conclude that no setting of θ will satisfy the designer's objective.

Now, we will show that when the number of samples is large enough, every Nash equilibrium of Γ_n is close to some Nash equilibrium of the underlying game. This result will lead us to consider convergence of optimizers based on empirical data to actual optimal mechanism parameter settings.

We first note that the function $\epsilon(s)$ is continuous in a finite game.

LEMMA 5. *Let S be a mixed strategy set defined on a finite game. Then $\epsilon : S \to \mathbb{R}$ is continuous.*

For the exposition that follows, we need a bit of additional notation. First, let (Z, d) be a metric space, and $X, Y \subset Z$ and define *directed Hausdorff distance* from X to Y to be

$$h(X, Y) = \sup_{x \in X} \inf_{y \in Y} d(x, y).$$

Observe that $U \subset X \Rightarrow h(U, Y) \leq h(X, Y)$. Further, define $B_S(x, \delta)$ to be an open ball in $S \subset Z$ with center $x \in S$ and radius δ. Now, let $\mathcal{N}_n$ denote all Nash equilibria of the game Γ_n and let

$$\mathcal{N}_\delta = \bigcup_{x \in \mathcal{N}} B_S(x, \delta),$$

that is, the union of open balls of radius δ with centers at Nash equilibria of Γ. Note that $h(\mathcal{N}_\delta, \mathcal{N}) = \delta$.

We can then prove the following general result.

THEOREM 6. *Suppose $|I| < \infty$ and $|A| < \infty$. Then almost surely $h(\mathcal{N}_n, \mathcal{N})$ converges to 0.*

We will now show that in the special case when Θ and A are finite and each Γ_θ has a unique Nash equilibrium, the estimates $\hat{\theta}$ of optimal designer parameter converge to an actual optimizer almost surely.

Let $\hat{\theta} = \arg\max_{\theta \in \Theta} W(\mathcal{N}_n(\theta), \theta)$, where n is the number of times each pure profile was sampled in Γ_θ for every θ, and let $\theta^* = \arg\max_{\theta \in \Theta} W(\mathcal{N}(\theta), \theta)$.

THEOREM 7. *Suppose $|\mathcal{N}(\theta)| = 1$ for all $\theta \in \Theta$ and suppose that Θ and A are finite. Let $W(s, \theta)$ be continuous at the unique $s^*(\theta) \in \mathcal{N}(\theta)$ for each $\theta \in \Theta$. Then $\hat{\theta}$ is a consistent estimator of θ^* if $W(\mathcal{N}(\theta), \theta)$ is defined as a supremum, infimum, or expectation over the set of Nash equilibria. In fact, $\hat{\theta} \to \theta^*$ a.s. in each of these cases.*

The shortcoming of the above result is that, within our framework, the designer has no way of knowing or ensuring that Γ_θ do, indeed, have unique equilibria. However, it does lend some theoretical justification for pursuing design in this manner, and, perhaps, will serve as a guide for more general results in the future.

6. RELATED WORK

The mechanism design literature in Economics has typically explored existence of a mechanism that implements a social choice function in equilibrium [10]. Additionally, there is an extensive literature on optimal auction design [10], of which the work by Roger Myerson [11] is, perhaps, the most relevant. In much of this work, analytical results are presented with respect to specific utility functions and accounting for constraints such as incentive compatibility and individual rationality.

Several related approaches to search for the best mechanism exist in the Computer Science literature. Conitzer and Sandholm [6] developed a search algorithm when all the relevant game parameters are common knowledge. When payoff functions of players are unknown, a search using simulations has been explored as an alternative. One approach in that direction, taken in [4] and [15], is to co-evolve the mechanism parameter and agent strategies, using some notion of social utility and agent payoffs as fitness criteria. An alternative to co-evolution explored in [16] was to optimize a well-defined welfare function of the designer using genetic programming. In this work the authors used a common learning strategy for all agents and defined an outcome of a game induced by a mechanism parameter as the outcome of joint agent learning. Most recently, Phelps et al. [14] compared two mechanisms based on expected social utility with expectation taken over an empirical distribution of equilibria in games defined by heuristic strategies, as in [18].

7. CONCLUSION

In this work we spent considerable effort developing general tactics for empirical mechanism design. We defined a formal game-theoretic model of interaction between the designer and the participants of the mechanism as a two-stage game. We also described in some generality the methods for estimating a sample Nash equilibrium function when the data is extremely scarce, or a Nash equilibrium correspondence when more data is available. Our techniques are designed specifically to deal with problems in which both the mechanism parameter space and the agent strategy sets are infinite and only a relatively small data set can be acquired.

A difficult design issue in the TAC/SCM game which the TAC community has been eager to address provides us with a setting to test our methods. In applying empirical game analysis to the problem at hand, we are fully aware that our results are inherently inexact. Thus, we concentrate on collecting evidence about the structure of the Nash equilibrium correspondence. In the end, we can try to provide enough evidence to either prescribe a parameter setting, or suggest that no setting is possible that will satisfy the designer. In the case of TAC/SCM, our evidence suggests quite strongly that storage cost could not have been effectively adjusted in the 2004 tournament to curb excessive day-0 procurement without detrimental effects on overall profitability. The success of our analysis in this extremely complex environment with high simulation costs makes us optimistic that our methods can provide guidance in making mechanism design decisions in other challenging domains. The theoretical results confirm some intuitions behind the empirical mechanism design methods we have introduced, and increases our confidence that our framework can be effective in estimating the best mechanism parameter choice in relatively general settings.

Acknowledgments

We thank Terence Kelly, Matthew Rudary, and Satinder Singh for helpful comments on earlier drafts of this work. This work was supported in part by NSF grant IIS-0205435 and the DARPA REAL strategic reasoning program.

8. REFERENCES

[1] R. Arunachalam and N. M. Sadeh. The supply chain trading agent competition. *Electronic Commerce Research and Applications*, 4:63–81, 2005.

[2] M. Benisch, A. Greenwald, V. Naroditskiy, and M. Tschantz. A stochastic programming approach to scheduling in TAC SCM. In *Fifth ACM Conference on Electronic Commerce*, pages 152–159, New York, 2004.

[3] Y.-P. Chang and W.-T. Huang. Generalized confidence intervals for the largest value of some functions of parameters under normality. *Statistica Sinica*, 10:1369–1383, 2000.

[4] D. Cliff. Evolution of market mechanism through a continuous space of auction-types. In *Congress on Evolutionary Computation*, 2002.

[5] D. A. Cohn, Z. Ghahramani, and M. I. Jordan. Active learning with statistical models. *Journal of Artificial*

Intelligence Research, 4:129–145, 1996.

[6] V. Conitzer and T. Sandholm. An algorithm for automatically designing deterministic mechanisms without payments. In *Third International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 128–135, 2004.

[7] D. Friedman. Evolutionary games in economics. *Econometrica*, 59(3):637–666, May 1991.

[8] R. Keener. *Statistical Theory: A Medley of Core Topics*. University of Michigan Department of Statistics, 2004.

[9] C. Kiekintveld, Y. Vorobeychik, and M. P. Wellman. An analysis of the 2004 supply chain management trading agent competition. In *IJCAI-05 Workshop on Trading Agent Design and Analysis*, Edinburgh, 2005.

[10] A. Mas-Colell, M. Whinston, and J. Green. *Microeconomic Theory*. Oxford University Press, 1995.

[11] R. B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, February 1981.

[12] S. Olafsson and J. Kim. Simulation optimization. In E. Yucesan, C.-H. Chen, J. Snowdon, and J. Charnes, editors, *2002 Winter Simulation Conference*, 2002.

[13] D. Pardoe and P. Stone. TacTex-03: A supply chain management agent. *SIGecom Exchanges*, 4(3):19–28, 2004.

[14] S. Phelps, S. Parsons, and P. McBurney. Automated agents versus virtual humans: an evolutionary game theoretic comparison of two double-auction market designs. In *Workshop on Agent Mediated Electronic Commerce VI*, 2004.

[15] S. Phelps, S. Parsons, P. McBurney, and E. Sklar. Co-evolution of auction mechanisms and trading strategies: towards a novel approach to microeconomic design. In *ECOMAS 2002 Workshop*, 2002.

[16] S. Phelps, S. Parsons, E. Sklar, and P. McBurney. Using genetic programming to optimise pricing rules for a double-auction market. In *Workshop on Agents for Electronic Commerce*, 2003.

[17] Y. Vorobeychik, M. P. Wellman, and S. Singh. Learning payoff functions in infinite games. In *Nineteenth International Joint Conference on Artificial Intelligence*, pages 977–982, 2005.

[18] W. E. Walsh, R. Das, G. Tesauro, and J. O. Kephart. Analyzing complex strategic interactions in multi-agent systems. In *AAAI-02 Workshop on Game Theoretic and Decision Theoretic Agents*, 2002.

[19] M. P. Wellman, J. J. Estelle, S. Singh, Y. Vorobeychik, C. Kiekintveld, and V. Soni. Strategic interactions in a supply chain game. *Computational Intelligence*, 21(1):1–26, February 2005.

APPENDIX

A. PROOFS

A.1 Proof of Proposition 1

$$\Pr\left(\max_{i\in I}\max_{b\in A_i\setminus a_i} u_i(b,a_{-i}) - u_i(a) \le \epsilon\right) =$$
$$= \prod_{i\in I} E_{u_i(a)}\left[\Pr(\max_{b\in A_i\setminus a_i} u_i(b,a_{-i}) - u_i(a) \le \epsilon | u_i(a))\right] =$$
$$= \prod_{i\in I}\int_{\mathbb{R}}\prod_{b\in A_i\setminus a_i} \Pr(u_i(b,a_{-i}) \le u+\epsilon) f_{u_i(a)}(u)du.$$

A.2 Proof of Proposition 2

First, let us suppose that some function, $f(x)$ defined on $[a_i, b_i]$, satisfy Lipschitz condition on $(a_i, b_i]$ with Lipschitz constant A_i. Then the following claim holds:

Claim: $\inf_{x\in(a_i,b_i]} f(x) \ge 0.5(f(a_i) + f(b_i) - A_i(b_i - a_i)$.

To prove this claim, note that the intersection of lines at $f(a_i)$ and $f(b_i)$ with slopes $-A_i$ and A_i respectively will determine the lower bound on the minimum of $f(x)$ on $[a_i, b_i]$ (which is a lower bound on infimum of $f(x)$ on $(a_i, b_j]$). The line at $f(a_i)$ is determined by $f(a_i) = -A_i a_i + c_L$ and the line at $f(b_i)$ is determined by $f(b_i) = A_i b_i + c_R$. Thus, the intercepts are $c_L = f(a_i) + A_i a_i$ and $c_R = f(b_i) + A_i b_i$ respectively. Let x^* be the point at which these lines intersect. Then,

$$x^* = -\frac{f(x^*) - c_R}{A} = \frac{f(x^*) - c_L}{A}.$$

By substituting the expressions for c_R and c_L, we get the desired result.

Now, subadditivity gives us

$$\Pr\{\bigvee_{\theta\in\Theta} \sup\{\phi^*(\theta)\} \le \alpha\} \le \sum_{j=1}^{5} \Pr\{\bigvee_{\theta\in\Theta_j} \sup\{\phi^*(\theta)\} \le \alpha\},$$

and, by the claim,

$$\Pr\{\bigvee_{\theta\in\Theta_j} \sup\{\phi^*(\theta)\} \le \alpha\} =$$
$$1 - \Pr\{\inf_{\theta\in\Theta_j} \sup\{\phi^*(\theta)\} > \alpha\} \le$$
$$\Pr\{\sup\{\phi^*(a_j)\} + \sup\{\phi^*(b_j)\} \le 2\alpha + A_j(b_j - a_j)\}.$$

Since we have a finite number of points in the data set for each θ, we can obtain the following expression:

$$\Pr\{\sup\{\phi^*(a_j)\} + \sup\{\phi^*(b_j)\} \le c_j\} =_{\mathcal{D}}$$
$$\sum_{y,z\in D: y+z\le c_j} \Pr\{\sup\{\phi^*(b_j)\} = y\} \Pr\{\sup\{\phi^*(a_j)\} = z\}.$$

We can now restrict attention to deriving an upper bound on $\Pr\{\sup\{\phi^*(\theta)\} = y\}$ for a fixed θ. To do this, observe that

$$\Pr\{\sup\{\phi^*(\theta)\} = y\} \le_{\mathcal{D}} \Pr\{\bigvee_{a\in\mathcal{D}:\phi(a)=y} \epsilon(a) = 0\} \le$$
$$\sum_{a\in\mathcal{D}:\phi(a)=y} \Pr\{\epsilon(a) = 0\}$$

by subadditivity and the fact that a profile a is a Nash equilibrium if and only if $\epsilon(a) = 0$.

Putting everything together yields the desired result.

A.3 Proof of Theorem 3

First, we will need the following fact:

Claim: Given a function $f_i(x)$ and a set X, $|\max_{x\in X} f_1(x) - \max_{x\in X} f_2(x)| \le \max_{x\in X} |f_1(x) - f_2(x)|$.

To prove this claim, observe that

$$|\max_{x\in X} f_1(x) - \max_{x\in X} f_2(x)| =$$
$$\begin{cases} \max_x f_1(x) - \max_x f_2(x) & \text{if } \max_x f_1(x) \ge \max_x f_2(x) \\ \max_x f_2(x) - \max_x f_1(x) & \text{if } \max_x f_2(x) \ge \max_x f_1(x) \end{cases}$$

In the first case,

$$\max_{x\in X} f_1(x) - \max_{x\in X} f_2(x) \le \max_{x\in X}(f_1(x) - f_2(x)) \le$$
$$\le \max_{x\in X} |f_1(x) - f_2(x)|.$$

Similarly, in the second case,

$$\max_{x\in X} f_2(x) - \max_{x\in X} f_1(x) \le \max_{x\in X}(f_2(x) - f_1(x)) \le \\ \le \max_{x\in X} |f_2(x) - f_1(x)| = \max_{x\in X} |f_1(x) - f_2(x)|.$$

Thus, the claim holds.

By the Strong Law of Large Numbers, $u_{n,i}(a) \to u_i(a)$ a.s. for all $i \in I, a \in A$. That is,

$$\Pr\{\lim_{n\to\infty} u_{n,i}(a) = u_i(a)\} = 1,$$

or, equivalently [8], for any $\alpha > 0$ and $\delta > 0$, there is $M(i,a) > 0$ such that

$$\Pr\{\sup_{n\ge M(i,a)} |u_{n,i}(a) - u_i(a)| < \frac{\delta}{2|A|}\} \ge 1 - \alpha.$$

By taking $M = \max_{i\in I} \max_{a\in A} M(i,a)$, we have

$$\Pr\{\max_{i\in I} \max_{a\in A} \sup_{n\ge M} |u_{n,i}(a) - u_i(a)| < \frac{\delta}{2|A|}\} \ge 1 - \alpha.$$

Thus, by the claim, for any $n \ge M$,

$$\sup_{n\ge M} |\epsilon_n(s) - \epsilon(s)| \le$$
$$\max_{i\in I} \max_{a_i\in A_i} \sup_{n\ge M} |u_{n,i}(a_i, s_{-i}) - u_i(a_i, s_{-i})| +$$
$$+ \sup_{n\ge M} \max_{i\in I} |u_{n,i}(s) - u_i(s)| \le$$
$$\max_{i\in I} \max_{a_i\in A_i} \sum_{b\in A_{-i}} \sup_{n\ge M} |u_{n,i}(a_i, b) - u_i(a_i, b)| s_{-i}(b) +$$
$$+ \max_{i\in I} \sum_{b\in A} \sup_{n\ge M} |u_{n,i}(b) - u_i(b)| s(b) \le$$
$$\max_{i\in I} \max_{a_i\in A_i} \sum_{b\in A_{-i}} \sup_{n\ge M} |u_{n,i}(a_i, b) - u_i(a_i, b)| +$$
$$+ \max_{i\in I} \sum_{b\in A} \sup_{n\ge M} |u_{n,i}(b) - u_i(b)| <$$
$$\max_{i\in I} \max_{a_i\in A_i} \sum_{b\in A_{-i}} (\frac{\delta}{2|A|}) + \max_{i\in I} \sum_{b\in A} (\frac{\delta}{2|A|}) \le \delta$$

with probability at least $1 - \alpha$. Note that since $s_{-i}(a)$ and $s(a)$ are bounded between 0 and 1, we were able to drop them from the expressions above to obtain a bound that will be valid independent of the particular choice of s. Furthermore, since the above result can be obtained for an arbitrary $\alpha > 0$ and $\delta > 0$, we have $\Pr\{\lim_{n\to\infty} \epsilon_n(s) = \epsilon(s)\} = 1$ uniformly on S.

A.4 Proof of Lemma 5

We prove the result using uniform continuity of $u_i(s)$ and preservation of continuity under maximum.

Claim: A function $f : \mathbb{R}^k \to \mathbb{R}$ defined by $f(t) = \sum_{i=1}^k z_i t_i$, where z_i are constants in $\mathbb{R}$, is uniformly continuous in t.

The claim follows because $|f(t) - f(t')| = |\sum_{i=1}^k z_i(t_i - t'_i)| \le \sum_{i=1}^k |z_i||t_i - t'_i|$. An immediate result of this for our purposes is that $u_i(s)$ is uniformly continuous in s and $u_i(a_i, s_{-i})$ is uniformly continuous in s_{-i}.

Claim: Let $f(a, b)$ be uniformly continuous in $b \in B$ for every $a \in A$, with $|A| < \infty$. Then $V(b) = \max_{a\in A} f(a, b)$ is uniformly continuous in b.

To show this, take $\gamma > 0$ and let $b, b' \in B$ such that $\|b - b'\| < \delta(a) \Rightarrow |f(a, b) - f(a, b')| < \gamma$. Now take $\delta = \min_{a\in A} \delta(a)$. Then, whenever $\|b - b'\| < \delta$,

$$|V(b) - V(b')| = |\max_{a\in A} f(a, b) - \max_{a\in A} f(a, b')| \le \\ \max_{a\in A} |f(a, b) - f(a, b')| < \gamma.$$

Now, recall that $\epsilon(s) = \max_i[\max_{a_i\in A_i} u_i(a_i, s_{-i}) - u_i(s)]$. By the claims above, $\max_{a_i\in A_i} u_i(a_i, s_{-i})$ is uniformly continuous in s_{-i} and $u_i(s)$ is uniformly continuous in s. Since the difference of two uniformly continuous functions is uniformly continuous, and since this continuity is preserved under maximum by our second claim, we have the desired result.

A.5 Proof of Theorem 6

Choose $\delta > 0$. First, we need to ascertain that the following claim holds:

Claim: $\bar{\epsilon} = \min_{s\in S\setminus\mathcal{N}_\delta} \epsilon(s)$ exists and $\bar{\epsilon} > 0$.

Since $\mathcal{N}_\delta$ is an open subset of compact S, it follows that $S \setminus \mathcal{N}_\delta$ is compact. As we had also proved in Lemma 5 that $\epsilon(s)$ is continuous, existence follows from the Weierstrass theorem. That $\bar{\epsilon} > 0$ is clear since $\epsilon(s) = 0$ if and only if s is a Nash equilibrium of Γ.

Now, by Theorem 3, for any $\alpha > 0$ there is M such that

$$\Pr\{\sup_{n\ge M} \sup_{s\in S} |\epsilon_n(s) - \epsilon(s)| < \bar{\epsilon}\} \ge 1 - \alpha.$$

Consequently, for any $\delta > 0$,

$$\Pr\{\sup_{n\ge M} h(\mathcal{N}_n, \mathcal{N}_\delta) < \delta\} \ge \Pr\{\forall n \ge M\ \mathcal{N}_n \subset \mathcal{N}_\delta\} \ge$$
$$\Pr\{\sup_{n\ge M} \sup_{s\in\mathcal{N}} \epsilon(s) < \bar{\epsilon}\} \ge$$
$$\Pr\{\sup_{n\ge M} \sup_{s\in S} |\epsilon_n(s) - \epsilon(s)| < \bar{\epsilon}\} \ge 1 - \alpha.$$

Since this holds for an arbitrary $\alpha > 0$ and $\delta > 0$, the desired result follows.

A.6 Proof of Theorem 7

Fix θ and choose $\delta > 0$. Since $W(s, \theta)$ is continuous at $s^*(\theta)$, given $\epsilon > 0$ there is $\delta > 0$ such that for every s' that is within δ of $s^*(\theta)$, $|W(s', \theta) - W(s^*(\theta), \theta)| < \epsilon$. By Theorem 6, we can find $M(\theta)$ large enough such that all $s' \in \mathcal{N}_n$ are within δ of $s^*(\theta)$ for all $n \ge M(\theta)$ with probability 1. Consequently, for any $\epsilon > 0$ we can find $M(\theta)$ large enough such that with probability 1 we have $\sup_{n\ge M(\theta)} \sup_{s'\in\mathcal{N}_n} |W(s', \theta) - W(s^*(\theta), \theta)| < \epsilon$.

Let us assume without loss of generality that there is a unique optimal choice of θ. Now, since the set Θ is finite, there is also the "second-best" choice of θ (if there is only one $\theta \in \Theta$ this discussion is moot anyway):

$$\theta^{**} = \arg\max_{\Theta\setminus\theta^*} W(s^*(\theta), \theta).$$

Suppose w.l.o.g. that θ^{**} is also unique and let

$$\Delta = W(s^*(\theta^*), \theta^*) - W(s^*(\theta^{**}), \theta^{**}).$$

Then if we let $\epsilon < \Delta/2$ and let $M = \max_{\theta\in\Theta} M(\theta)$, where each $M(\theta)$ is large enough such that $\sup_{n\ge M(\theta)} \sup_{s'\in\mathcal{N}_n} |W(s', \theta) - W(s^*(\theta), \theta)| < \epsilon$ a.s., the optimal choice of θ based on *any* empirical equilibrium will be θ^* with probability 1. Thus, in particular, given any probability distribution over empirical equilibria, the best choice of θ will be θ^* with probability 1 (similarly, if we take supremum or infimum of $W(\mathcal{N}_n(\theta), \theta)$ over the set of empirical equilibria in constructing the objective function).