Recombinant Estimation for Normal-Form Games, with Applications to Auctions and Bargaining

Charles H. Mullin*
David H. Reiley†*

This Version: November 2004
First Version: January 1999

Abstract

When analyzing economic games, researchers frequently estimate quantities describing group outcomes, such as the expected revenue in an auction. For such applications, we propose an improved statistical estimation technique called "recombinant estimation." The technique takes observations of players' strategies and recombines them to compute all possible group outcomes that could have resulted from different player combinations. In applications to an auction and a bargaining game, the improved efficiency of our estimator is equivalent to increasing the sample size between 25% and 200%. We discuss how to design experiments in order to utilize recombinant estimation. We also discuss practical computational issues, and provide software for computing estimates and standard errors.

JEL Classification: C70, C13 and C92.

Keywords: Simultaneous-Move Games, Experimental Statistics

* Mullin can be contacted at charlie.mullin@bateswhite.com, and Reiley can be reached at reiley@eller.arizona.edu. We thank J.S. Butler, Yan Chen, Rachel Croson, Rosemarie Nagel, Karim Sadrieh, and Mark Satterthwaite for valuable discussions. Mike Urbancic has done a wonderful job writing flexible, user-friendly software in Microsoft Excel, so that others may easily use our estimator without tedious programming. We are also grateful for the comments of an anonymous referee and an associate editor.

† Reiley gratefully acknowledges the National Science Foundation for support through grants SBR-9811273 and SES-0094800.

1. Introduction

Economists frequently study behavior in simultaneous-move games with multiple players, from auctions to Bertrand oligopoly to public-good provision. Recent interest in behavioral game theory has generated a wealth of experimental data on individuals' play in a wide variety of such games. These data can be used to make inferences about various quantities of economic interest, such as auction revenues, market prices, and levels of public-good provision. Often, experiments on multi-player games group each subject randomly and arbitrarily with anonymous opponents. In such simultaneous-move games with arbitrary groupings, it seems reasonable to assume that the identities of any given player's rivals do not alter her strategy. This assumption implies that the hypothetical game outcomes which would result from grouping a subject with players she did not actually face are just as valid as the game outcome resulting from the group of players she did play against. In this paper, we develop an estimation technique, which we call "recombinant estimation," designed to extract maximal information from observed strategies in order to estimate various properties of simultaneous-move games.

For concreteness, we present the example that originally motivated this paper. List and Lucking-Reiley (2000) compare two different sealed-bid auction formats. They conducted 15 identical two-person, uniform-price auctions for identical pairs of goods. This yielded data from 30 bidders, all of whom faced the identical bidding situation. They repeated this procedure with 30 additional bidders in 15 different multi-unit Vickrey auctions. To compare the two auction formats, they conducted *t*-tests, which involve estimating both the mean and the standard error of the mean for each auction format. Recall that with n independent observations, the standard error of the mean equals the standard deviation of the data divided by the square root of n. In order to test whether the mean bids were equal across auction formats, they computed the mean bid, and computed the standard error as the standard deviation divided by the square root of 30. They also compared revenues across auction formats, but since there were only 15 revenue observations (15 auctions with 2 bidders each), the standard error for this case was equal to that standard deviation divided only by the square root of 15. This seemed somehow unfair, as the same amount of information led to relatively higher standard errors for revenues than for bids.

This observation led us to realize that in a sealed-bid auction, the pairings of bidders were arbitrary, so we could observe the revenues that would have resulted from all 435 possible pairings (the number of combinations of 30 bidders taken two at a time). Computing the mean over these 435 different revenue observation clearly gives a more precise estimate of the mean revenue in the population of auctions with all possible bidders, since it uses more information. But what should be the standard error of the mean? Surely one can't just divide the standard deviation by the square root of 435; there's not *that* much information available! In this paper, we formalize the intuition just discussed, presenting a recombinant estimator for use in a variety of normal-form game applications, and deriving correct standard-error estimates so that the estimator may be used for inference purposes.

We are not the first to propose recombining players to attain more-accurate estimators. For example, in their analysis of a bargaining game, Mitzkewitz and Nagel (1993) use the complete contingent strategies for each player to compute recombinant estimates of expected profits. Similarly, Mehta *et al.* (1994) utilize recombinant estimation to determine the frequency of success in pure coordination games. Verifying the intuition of this approach, Friedman (1996) shows that players learn a game faster if they receive the expected payoff across all possible opponents, as opposed to the payoff based solely on the actions of one randomly chosen opponent.

What is new in this paper is the placement of the idea of recombinant estimation into a formal econometric framework. First and foremost, we derive standard errors for this estimation technique, which both quantifies the efficiency improvement associated with this technique and enables hypothesis testing. For example, in the motivating example of List and Lucking-Reiley (2000), we attain efficiency improvements of 19 to 68 percent, depending on the quantity estimated, which is roughly equivalent to increasing the sample size by 25 to 200 percent. When applied to the Mitzkewitz and Nagel (1993) data, the recombinant estimator improves efficiency 16 to 68 percent, with the efficiency gains increasing with the complexity of the strategy space.

Second, formalizing the estimation technique clearly demonstrates when it can and cannot be used, which translates into some recommendations for experimental design. The key is to make sure that players face the same public environment in each session of the experiment (e.g., the number of players, the structure of the payoff function, the distribution of information, etc.), so that the recombination across sessions is meaningful. In other words, the only items allowed to change across sessions are the actual players taking each role and the realization of any information that is private to a particular player.

Third, recombinant estimation can result in a combinatorial explosion in the number of computations required. Depending on the number of players per group, a few dozen actual observations might spawn billions of hypothetical observations. We show that one can achieve almost all of the efficiency gains of recombinant estimation by considering only a random sample of several thousand of the hypothetical observations, an estimation task readily achieved by modern computers in seconds or less. Similar practical considerations apply to the estimation of standard errors for the estimator, where the combinatorial explosions are even more rapid than in the computation of the estimator itself. Using these practical insights, we have implemented our estimator in user-friendly computer software, available for download from the second author's Web site.[1]

[1] The software is written with macros in Microsoft Excel. To use it, one copies one's data into the Excel workbook, enters the group outcome formula, and executes the software to obtain estimates of the mean group outcome and the standard error of this mean. At this writing, the software works only for symmetric games, but we hope in future to

Finally, recombinant estimation bears some similarities to the statistical technique of bootstrapping, so it may be worth pointing out the differences. A traditional bootstrap reuses the data in a similar manner, but they do so in order to attain more-robust confidence intervals for the test statistic of interest. More specifically, bootstrapping provides a two-step procedure to estimate numerically the sampling distribution of an estimator. The bootstrap first estimates the data-generating process from the empirical distribution of the data, and then uses Monte Carlo simulations based on this estimated data-generating process to estimate the sampling distribution of the estimator. Thus, while bootstrapping is a method for doing inference with an existing estimator, recombinant estimation constitutes a new, more efficient estimator for a population parameter.[2] In other words, we switch the focus from inference about a given estimator to developing a more efficient estimator.

The remainder of the paper is organized as follows: Section 2 develops the estimator and its distribution. Section 3 addresses the efficiency of the estimator. Section 4 provides two empirical examples, while Section 5 discusses how to design experiments in order to take advantage of this estimation technique. Section 6 presents practical computational considerations, and Section 7 concludes.

2. Estimators for Game Outcomes

In this section, we develop the recombinant estimator for the expected value of an outcome in a simultaneous-move game. For the sake of exposition, we start with relatively simple games, and add more generality in later subsections. In each case we consider, the estimator is a sample average that satisfies a central limit theorem, so it has an asymptotically normal sampling distribution.

---

provide software for asymmetric games as well. The software is currently available at <http://eller.arizona.edu/~reiley/papers/RecombinantGames.html>.

[2] Similarly, the Fisher randomization test is solely a method for doing inference, not an estimation technique. See Moir (1998) for details.

Before developing the estimator in detail, we introduce some notation. Assume that the econometrician observes n groups of k players each, and let $m = nk$ denote the total number of players. Let x_i be the strategy observed for the ith player and p_j be the vector of players who participated in group j. For concreteness, we might think of the first group of players as $p_1 = (1, 2, \ldots, k)$, $p_2 = (k + 1, k + 2, \ldots, 2k)$ as the second group, and so on. Denote p_{jt} as the tth player in group j. Define y_j as the outcome which results from the strategies played by the set of players p_j, and let g be the function which maps strategies to group outcomes, so that $y_j = g\left(x_{p_{j1}}, x_{p_{j2}}, \ldots, x_{p_{jk}}\right)$.

The goal is to estimate the expected value of y over the entire population of possible players. Let y have a mean of μ and a variance of σ^2. For comparison to the recombinant estimator, we first present the baseline estimator most commonly observed in practice. This baseline estimator computes the simple average of the group outcome for each of the n observed groups. We denote this baseline estimator by:

$$\hat{\mu} = \bar{y} = (1/n)\sum_{j=1}^{n} y_j \tag{2-1}$$

where j indexes the n groups. The variance of this estimator is

$$\operatorname{var}(\bar{y}) = \sigma^2 / n \,. \tag{2-2}$$

This estimator serves as our reference point for efficiency comparisons to other proposed estimators.

In the following subsections, we first develop our recombinant estimator for symmetric two-player games and then generalize it to symmetric k-player games. We next develop a recombinant estimator for fully asymmetric k-player games, and finally consider general k-player games. The special cases provide a simplified setting in which to explain the intuition behind the recombinant estimation strategy.

2.1 Symmetric Two-Player Games

Suppose we observe four players matched in two pairs in a two-player, symmetric game. Let y_1 represent the outcome for players 1 and 2, and y_2 represent the outcome for players 3 and 4. Using the baseline estimator, the average of these two outcomes would be our estimate. However, one might be able to improve upon this estimator by considering alternative combinations of the players. In particular, the baseline estimator ignores the potential pairings {{1,3},{1,4},{2,3},{2,4}}.

The recombinant estimator considers these potential pairings. Recall that p_j represents the set of players in group *j*. For the baseline estimator, *j* indexed only those groups actually observed, so it ranged from 1 to *n*. Now we let P represent the collection of all possible 2-player sets which can be taken from the *m* players, and let *j* index all the elements of P, so that *j* now ranges from 1 to $\mathrm{J} \equiv m!/(2!(m-2)!)$. The recombinant estimator is the simple average of these J combinations:

$$\hat{\mu} = \left(\frac{1}{\mathrm{J}}\right)\sum_{j=1}^{\mathrm{J}} y_j \ . \tag{2-3}$$

Note that the set P fails to distinguish between the pairings $\{i, j\}$ and $\{j, i\}$. Due to the symmetry in the outcome function, these two outcomes correlate perfectly, so the second adds no new information once the first has been considered.

We have chosen to ignore the hypothetical outcomes in which a player is matched with herself. In some applications, the researcher may find it appropriate to include additional terms of the form $\{i, i\}$ in the estimator of the mean. The additional information would improve the efficiency of the estimator in small samples, but the effect is negligible in large samples, as the number of $\{i, i\}$ terms becomes negligible relative to the total number of $\{i, j\}$ terms. The two approaches are thus asymptotically equivalent, but the present form is notationally less burdensome. It is relatively straightforward to adapt the equations in this paper to include these additional pairings.

To present the variance of this estimator requires some additional notation. Let $\varphi \equiv \text{cov}(y_i, y_j)$ when groups i and j share exactly one common player. Note that since each player's strategy is independent of all other players' strategies, when there is no intersection between two groups of players the covariance is zero. So,

$$\begin{aligned} \text{var}(\hat{\mu}) &= (1/\text{J})^2 \left[\sum_{j=1}^{\text{J}} \text{var}(y_j) + \sum_{j=1}^{\text{J}} \sum_{i \neq j} \text{cov}(y_j, y_i) \right] \\ &= (1/\text{J})^2 \left[\text{J}\,\sigma^2 + m(m-1)(m-2)\varphi \right] \\ &= \frac{2}{m(m-1)}\sigma^2 + \frac{4(m-2)}{m(m-1)}\varphi \end{aligned} \tag{2-4}$$

where $m(m-1)(m-2)$ is the number of covariance terms with one common player across the two groups. Recall that $m = 2n$. Therefore, when φ is positive, $n \cdot \text{var}(\hat{\mu})$ converges to 2φ.[3] From (2-2), $n \cdot \text{var}(\bar{y})$ converges to σ^2. Thus, the asymptotic-efficiency gain of the recombinant estimator relative to the baseline estimator is $2\varphi/\sigma^2$. Proposition 1 in the following section proves that the recombinant estimator is weakly more efficient than the baseline estimator.

When φ is zero, $n^2 \cdot \text{var}(\hat{\mu})$ converges to $\sigma^2/2$, which is a faster rate of convergence (n^2 instead of n) than the baseline estimator. In general, whenever the covariance is zero, regardless of the symmetry of the game or the number of players involved, the recombinant estimator converges at a faster rate than the baseline estimator. Although this case is interesting, it is atypical: for example, two hypothetical auctions with the same aggressive bidder will both tend to have high revenues, and the opposite will be true for two hypothetical auctions with the same conservative bidder. For that reason, from this point onward we concentrate on the case where φ is strictly positive.

[3] The average covariance across games with at least one shared player, φ, must be non-negative.

For inference purposes, it is important both to be able to estimate the variance given by (2-4) and to demonstrate that this variance fully describes the distribution of the estimator. First, to estimate the variance, one can use all the hypothetical outcomes y_j to produce consistent estimators of σ^2 and of φ, and then substitute the resultant estimates into (2-4). Since (2-4) gives the exact equation for the variance, this estimate of the variance differs from the true variance only insofar as the estimates of σ^2 and φ differ from their population values. We will present more details on computing standard errors in Section 6.2. For now, it is sufficient to note that the recombinant estimation technique can be used to estimate both σ^2 and φ, which increases the accuracy with which these parameters are estimated. Second, the central limit theorem establishes that the recombinant estimator follows a normal distribution. Furthermore, the estimator is always a sample average (of repeated draws from a stationary distribution) and these averages tend to converge to normal distributions relatively quickly.[4]

2.2 Symmetric *K*-Player Games

With more than two players in a symmetric game, there are many more recombinations possible. Instead of merely $m(m-1)/2$ different hypothetical outcomes from the m players, we now have the number of combinations of m objects taken k at a time. Now we let P represent the collection of all possible k-player sets which can be taken from the m players, and let j index all the elements of P, so that j now ranges from 1 to $\mathrm{J} \equiv m!/(k!(m-k)!)$. Under this new definition of P, the recombinant estimator is still given by (2-3), $\hat{\mu} = (1/J)\sum_{j=1}^{\mathrm{J}} y_j$.

To compute the asymptotic variance of this estimator, we again need to account for the fact that when two groups have at least one overlapping player their outcomes correlate. In principle, we might expect two groups with more than one overlapping player, such as {1,2,3}

[4] When the sample size is sufficiently small such that either σ^2 and φ are poorly estimated or the normal approximation fails, a bootstrap is unlikely to improve inference since it relies on similar asymptotic assumptions.

and {1,2,4}, to have a higher covariance than two groups with just a single overlapping player, such as {1,2,3} and {1,4,5}. However, it becomes cumbersome to keep track of the $k - 1$ different types of covariances, so we consider the average covariance between any two groups that overlap by at least one player:

$$\varphi \equiv E\{\mathrm{cov}(y_i, y_j) | \; p_i \cap p_j \neq \varnothing \text{ and } p_i \neq p_j\} \tag{2-5}$$

where the expectation is taken with respect to the relative frequency with which each possible combination of repeated players occurs.

Using this notation, the variance of the estimator is:

$$\begin{aligned} \mathrm{var}(\hat{\mu}) &= (1/\mathrm{J})^2 \left[\sum_{j=1}^{\mathrm{J}} \mathrm{var}(y_j) + \sum_{j=1}^{\mathrm{J}} \sum_{i \neq j} \mathrm{cov}(y_j, y_i) \right] \\ &= n^{-1} \left[A_n \sigma^2 + (1 - A_n) k\varphi \right] \end{aligned} \tag{2-6}$$

where $A_n = n/\mathrm{J}$. Recalling that J is of order n^k, it is relatively straightforward to demonstrate that $A_1 = 1$, $\partial A_{\mathrm{n}}/\partial n < 0$ and $\lim_{n \to \infty} A_n = 0$. Thus, $n \cdot \mathrm{var}(\hat{\mu})$ converges to $k\varphi/n$ and its asymptotic efficiency relative to the baseline estimator is $k\varphi/\sigma^2$. Note that the two-player game of the previous section is a special case of a k-player symmetric game for which $A_n = (2n - 1)^{-1}$.

2.3 Fully Asymmetric *K*-Player Games

The generalization from k–player symmetric games to k-player asymmetric games is reasonably straightforward. We assume that the game is fully asymmetric, in the sense that each of the k different players enters the outcome function in a unique way. We will relax this assumption in section 2.4.

Let S_i be the set of individuals who played in the ith player role. Then, the set of all possible groups of players becomes: $\mathrm{P} = \{(i_1, \ldots, i_k) | \; i_j \in S_j \; \forall j \in \{1, \ldots, k\}\}$. There are $\mathrm{J} = n^k$ elements in P. Under this new definition of P, (2-3) and (2-6) give the recombinant estimator and its variance, respectively. Thus, the efficiency of this estimator relative to the baseline estimator remains $k\varphi/\sigma^2$.

Note that treating a symmetric game as an asymmetric game, which reduces the number of allowable combinations, leaves the asymptotic efficiency of the estimator unchanged (the covariance φ remains unchanged). As illustrated in Section 6 on computational methods, this result occurs because the limit on the information from player i is the expected value of the outcome given that player i is in the game. As the number of recombinations considered increases, the estimate of this conditional expectation improves, but the number of such conditional expectations remains fixed. As the sample size gets large, the conditional expectation for each player is essentially known under either a symmetric or asymmetric treatment, leaving the standard error of the estimator to be determined by the variability across these conditional expectations.

2.4 General *K*-Player Games

Some games of interest have both symmetry and asymmetry between players. For example, in a sealed-bid double-auction call market with b buyers and s sellers, the buyers are all symmetric to each other (they each have the same strategy space), and the sellers are all symmetric to each other, but there is asymmetry between buyers and sellers. Such a game has symmetry within player roles, but asymmetry between player roles.

Let t be the number of player roles and t_i be the number of players of type i in each group, e.g., $t = 2$, $t_1 = b$ and $t_2 = s$ in the double-auction example above. As before, let S_i be the set of individuals who played in the ith player role, e.g., S_1 contains all of the buyers in the above example. Then, the set of all possible groups of players becomes:

$$\mathrm{P} = \left\{ \left(\left(i_{11}, \ldots, i_{1t_1}\right), \ldots, \left(i_{t1}, \ldots, i_{tt_t}\right) \right) \middle|\ i_{j*} \in S_j \ \forall j \in \{1, \ldots, t\} \right\}.$$

Under this new definition of P, (2-3) and (2-6) give the recombinant estimator and its variance, respectively. Again, the efficiency of this estimator relative to the baseline estimator is $k\varphi/\sigma^2$.

3. Theoretical Results on Recombinant Estimation

In this section, we present two propositions. The first proposition demonstrates that the recombinant estimator is weakly more efficient than the baseline estimator. The second proposition describes under what circumstances the recombinant estimator is strictly more efficient than the baseline estimator. All proofs are in Appendix A.

By proving that the average covariance between group outcomes which share overlapping strategies is less than one k^{th} of the variance of the group outcome function, Proposition 1 shows that the recombinant estimator is weakly more efficient than the baseline estimator.

Proposition 1: The recombinant estimator is weakly more efficient than the baseline estimator.

Furthermore, the proof of Proposition 1 illustrates under what conditions the recombinant estimator is strictly more efficient than the baseline estimator. In particular, it is sufficient that the outcome function is not an additively separable function of the strategies, i.e. $y_j = g\left(x_{p_{j1}}, \ldots, x_{p_{jk}}\right) \neq g_1\left(x_{p_{j1}}\right) + \ldots + g_k\left(x_{p_{jk}}\right)$.[5]

Proposition 2: The recombinant estimator is strictly more efficient than the baseline estimator if the outcome function is not additively separable in the strategies.

Proposition 2 indicates that recombinant estimation provides strict efficiency gains for many different group outcome functions. The only exception is where the outcome is a linearly separable function of the players' strategies. To illustrate an example of this exception, suppose one is interested in the overall economic efficiency (relative to the Pareto optimum) that can be

[5] If recombination across player roles is feasible, then asymmetry across player roles is also sufficient for the recombinant estimator to be strictly more efficient, i.e. g_i not equal to g_j for at least one pair of i and j.

obtained in a public-good-contribution game. The game might involve k players each endowed with 100 units of money, which they can either keep or contribute to the public good. When player i contributes x_i to the public good, that money gets doubled and then divided equally among the players for their payoffs. In this game, the x_i s are the strategies, and the group outcome is the total payoff to all players. The total payoff is equal to the total amount of public good provided plus the total amount of private good kept by the players, which turns out to be $100k + \Sigma x_i$. Since this is linear in the strategies, there is no econometric efficiency gain to be had. Intuitively, since the group outcome just involves summing up the individual outcomes, the recombinant estimator turns out to be numerically equivalent to the baseline estimator – it includes each term in the sum multiple times, but the average turns out to be identical. In the next section, we demonstrate two examples where the outcome function is nonlinear in the strategies, and compute the gains that occur.

4. Applied Examples

We present two examples of recombinant estimation. First, we estimate expected revenues and the allocation of goods in multi-unit auctions, from experiments by List and Lucking-Reiley (2000). Appendix B provides step-by-step details for the implementation of the recombinant estimator for this example. Second, we consider expected earnings of players in an ultimatum bargaining game, from experiments by Mitzkewitz and Nagel (1993).

4.1 Auctions

List and Lucking-Reiley (2000, henceforth LLR) compared two different sealed-bid auction formats in an experiment at a sportscard trading show. Each of the 328 different bidders participated in a single two-bidder, two-unit auction for sportscards, with the items' book values ranging from \$3 to \$70. Half the bidders bid in auctions with the uniform-price highest-rejected-bid price rule, while the other half bid in auctions with the generalized Vickrey price rule. Each auction offered a pair of identical sportscards, and each bidder had the opportunity to submit two bids (one for each of the units of the good). After collecting the bid sheets from individuals over a period of several hours, the auctioneer randomly paired up individuals from the same type of auction in order to compute the results. Bidders were asked to return to the same booth at an appointed time in order to learn the results of the auction and conduct the transaction for any cards they had won.

LLR had four main results. First, second-unit bids were lower in the uniform-price treatment than in the Vickrey treatment, as predicted by Nash equilibrium theory. Second, first-unit bids were higher in the uniform-price treatment than in the Vickrey treatment, an effect unpredicted by theory. Third, the allocations of goods were significantly different across auction formats with the uniform-price treatment resulting in more split allocations relative to the Vickrey treatment. Fourth, there was no statistically significant difference in revenues across auction formats. LLR's first two results involved analyses of the mean levels of bids. Since bids are individual strategies rather than group outcomes, recombinant estimation cannot improve econometric efficiency for mean bid levels. However, the last two results involve quantities (revenues and allocations) that are group outcomes. Since each bidder made his bids independently, we can use the recombination techniques of the present paper to provide improved estimates of the mean revenues and the frequency of split allocations for each auction format.

Tables 1 and 2 compare the baseline, non-recombinant estimates from LLR with recombinant estimates from the same data. Table 1 presents results for the proportion of split allocations under each format. There are results for ten different auction treatments: five Vickrey and five uniform-price (the five different treatments involved different goods and different types of bidders, as shown in the rows of the table). Moving from the baseline to the recombinant estimator, the point estimates change by moderate amounts either up or down, and the standard errors uniformly become smaller. The estimated variance of the recombinant estimator is 44 to 68 percent smaller than that of the baseline estimator, averaging a 57 percent improvement. This efficiency improvement substantially increases the power of *t*-tests of the hypothesis that Vickrey and uniform-price auctions produce equivalent allocations; for example, the two-tailed *p*-value for this test in the first treatment (Sanders 1989 card auctioned to Nondealers) decreases from 0.034 with the baseline estimator to 0.0001 with the recombinant estimator. This improvement in efficiency is approximately equivalent to a threefold increase in the sample size!

Table 2 presents similar results for auction revenues. Again, the standard errors decrease when moving from the baseline to the recombinant estimator. Overall, the estimated variance of the recombinant estimator is 26 percent smaller than that of the baseline estimator, which is about equivalent to a 35 percent increase in the sample size. In general, the difference in revenues across auction formats remains statistically insignificant, but the recombinant estimator provides marginal evidence (*p*-value of 0.06) that the Vickrey format produces lower revenues for the Jordan 1989 card.

4.2 Bargaining

Mitzkewitz and Nagel (1993) analyze the ultimatum game. Though a bargaining game does not typically involve simultaneous moves, Mitzkewitz and Nagel use the "strategy method" to collect data on complete contingent strategies, which converts the extensive-form, sequential-move bargaining game into a normal-form, simultaneous-move game. The basic idea of the ultimatum game is that Player A makes a take-it-or-leave-it offer to split a cash-valued "cake" with Player B. Player B either accepts the offer so that both players earn the indicated amounts, or rejects the offer and both players receive nothing. We consider the "demand form" of the game developed by Mitzkewitz and Nagel, in which the cake size is known to Player A but unknown to Player B. All Player B knows is the size of Player A's demand, but not the size of the cake or the relative fraction A has demanded. The specific rules of the game are as follows:

1) The cake may be one of six sizes: 1, 2, 3, 4, 5, or 6 Taler, a fictitious currency worth approximately 70 cents.
2) Player A commits to an offer to player B for each of the six possible cake sizes. Offers must be an element of the set $\{0, 0.5, 1.0, \ldots, 5.5, 6.0\}$ and cannot exceed the size of the cake.
3) Meanwhile, player B, with no knowledge of the cake size, decides which of the possible demands from zero to six she will accept, and which she will reject.
4) After both players have committed to their strategies, a die is rolled to determine the actual size of the cake. Given the cake size, player A's demand is revealed to player B, who either accepts or rejects it according to her previously elicited strategy. If accepted, player A receives what he demanded and player B receives the remainder of the cake. If rejected, both players receive zero.

This setting fits well the estimation technique of this paper. In fact, Mitzkewitz and Nagel themselves computed recombinant estimates of expected profits in their analysis (pg. 179). However, they did not provide standard errors of their estimates, nor did they evaluate the econometric efficiency of the procedure, so we extend their work in this direction. Table 3 contains the recombinant estimates,[6] their associated standard errors, and the relative efficiency gain of the recombinant estimator.[7]

This estimation produces two interesting results. First, recombinant estimation yields a 16 to 66 percent efficiency gain in estimating the expected profits of players A and B. The greater gains are realized at the higher cake sizes (where the strategy space is relatively rich). In general, the gains to recombinant estimation are roughly proportional to the variance in players'

[6] The difference in the point estimates presented here from those in Mitzkewitz and Nagel is attributable to the fact that Mitzkewitz and Nagel computed their estimates by hand.

[7] The experiment involved eight repeated trials with each player facing a different rival each round. Following Mizkewitz and Nagel, we recombine not just across individuals within a round, but also across these rounds. We correct the standard errors to account for the possibility of a positive correlation between strategies submitted in different rounds by the same bidder, which is distinct from the positive covariance already assumed to exist between observations using the strategy of the same player in the same round. For ease of exposition, we omit an explicit discussion of the equations used for this purpose, but they are a straightforward generalization of the equations already presented.

strategies. To be exact, the gain is proportional to the variance in the outcome function induced by the variance in players' strategies. Second, the estimated standard errors allow us to test an interesting hypothesis: whether the expected profit to player A is monotonically increasing in the cake size. Mitzkewitz and Nagel note that the estimated expected payoff to player A takes a maximum at cake size four and decreases from that point onward, so the point estimates look as if player A’s profits are not monotonic in the cake size. However, they were unable to compute a formal test of this hypothesis, because they did not know the distribution of their estimates. The estimated expected revenue decreases by approximately 0.128 from a cake of size four to one of size six. The *t*-statistic for this change is 1.88, so we are able to reject at the 5% significance level the hypothesis that type-A players’ profits are monotonically increasing in the cake size. Furthermore, the *t*-statistic associated with the baseline estimator is only 1.13, which would be insufficient evidence to reject the null hypothesis. Thus, in this example the recombinant estimation technique generates a formal rejection of the null hypothesis where standard estimation techniques did not.

5. Implications for Experimental Design

Our results show that recombinant estimation can produce large improvements in efficiency for estimating group outcomes from experimental data on auctions and other simultaneous-move games. We find three key implications of our work for experimental design. First, in dynamic games, eliciting complete contingent strategies maximizes the returns to recombinant estimation. Second, in order to recombine across groups of players, all group-level parameters must be held constant for each group, though individual-level parameters may vary. Third, in order to use recombinant estimation in an experiment with repeated trials, it is important to have an experimental design that minimizes repeated contact between the same players. Each of these implications is discussed in more detail below.

5.1 Complete Strategies

Consider a two-player, dynamic game in which player A moves first, then player B responds. Suppose player A has five potential moves labeled one through five. Consider a player in position A who chose option two. This player may only be recombined with players in position B whose opponent also chose option two. Alternatively, suppose the experimenter had complete contingent strategies from each player in position B. Then, each player in position A could be recombined with all players from position B. To attain complete contingent strategies and take full advantage of recombinant estimation, experimenters may wish to consider the use of the strategy method (as done in Mitzkewitz and Nagel, 1993).

5.2 Consistency Across Sessions

The experimenter needs to make sure that experimental conditions are held constant across experimental sessions in a way that allows for recombination, a technique known as a "yoked" design in experimental psychology. When using a yoked design, all sessions of the experiment will appear ex-ante identical to every player. Thus, switching a player from session *i* to session *j* will not alter her strategy. Typically changing an experimental design into a yoked design requires minimal effort. An example may be helpful.

Kagel *et al.* (1987) tested the theory of auctions with affiliated private values. In each experimental auction, they used the following procedure. First, they chose a locational parameter x_0 randomly from a uniform distribution on [\$25, \$125]. Second, they chose a set of six private values x_i for the six bidders, with each x_i drawn independently from the uniform distribution on $[x_0 - \varepsilon, x_0 + \varepsilon]$, where ε was a constant equal to either \$6, \$12, or \$24, depending on the round of the experiment. Each bidder learned his private value x_i, but did not necessarily learn the value of x_0, so he might remain uncertain about the distribution from which his value came. This generated "affiliation" between bidders' values, as described by Milgrom and Weber (1982).

That is, when a bidder learns that his value is high, he also learns that other bidders' values are likely to be relatively high.

To keep the discussion short, we consider only the first auction of their experiment, observed in each of seven separate sessions. If the design were yoked, there would be over 5.2 million different hypothetical revenue observations (combinations of 42 different subjects taken six at a time), instead of the mere seven observed outcomes. The exact amount of efficiency improvement would depend on the actual distribution of the data, but it is likely to be quite large. Unfortunately, the locational parameter x_0 makes it impossible to recombine bidders. The value of x_0 was drawn independently across sessions, leaving no two groups of six bidders with the same value. In order to use the recombinant estimator, one would have to hold the locational parameter fixed across sessions. Additionally, the range-of-values parameter ε, would need to be held constant across sessions. Note that it is not necessary to use the same set of individual values x_i in each experimental session, because the bids from bidders with different individual valuations could still be recombined, so long as the group-level parameters are the same. In fact, it would be preferable to generate different random bidder values across sessions, because this would create a better representation of the full population to be studied: all bidders who might possibly have bid in an auction with the particular values of x_0 and ε.

5.3 Learning

A key assumption underlying recombinant estimation is that strategies in different sessions are all independent draws from the same distribution. For example, in games with repeated play, in order to combine one individual's strategy in round 2 with another individual's strategy in round 10, one must believe that there is no systematic change in play over time. Although there exists a large literature on testing the hypothesis that two samples are drawn from the same distribution, the sample sizes in most experiments are sufficiently small such that the power of these tests would be very low. In lieu of focusing on these formal tests, we only consider two deviations from this assumption: learning about the game, and learning about one's opponents.

Suppose players learn about the game over successive rounds of play. For example, bidders gradually learn to bid lower. Under this scenario, bids in round 2 are drawn from a different distribution than bids in round 10. Thus, it is inappropriate to assume that a player in round 10 would have made the same bid if she were returned to round 2. In other words, one should not recombine across rounds, but only across sessions within the same round.

Another complication is the possibility of learning not just about the game but also about one's rival players: for example, one group of players learns to submit high bids because they played against a rival who bids unusually high in early rounds. If subjects do learn about their specific group of rivals in repeated play, then recombining across groups in any round after the first would be inappropriate. Although it might be possible to construct a recombinant estimator that also controls for group effects of this type, this task is beyond the scope of the present paper.

In order to avoid learning about one's rivals in an experiment with repeated trials, the experimenter might wish to choose an experimental design that minimizes the possibility for this type of learning. For example, Mitzkewitz and Nagel's experimental design ensured that over the eight rounds of the experiment, no player ever played against the same rival twice. Thus, in section 4.2, we feel relatively sanguine about recombining within rounds for that experiment. An even stronger insurance would be the "zipper design" sometimes used in experiments, where in repeated rounds a player never encounters anyone who might have been influenced by that player previously. (That is, one never plays one's previous rivals again, nor one's rivals' rivals, nor one's rivals' rivals' rivals, and so on.)

6. Computational Methods

Even a relatively small number of observations (seven groups of six subjects each) can yield a very large number of combinations. This explosive growth in the number of combinations, both with respect to the number of players in a game and the total number of players, could become computationally burdensome. Fortunately, the vast majority of the efficiency gain from recombinant estimation can be obtained from a relatively small fraction of those possible combinations. We show below that 100 combinations per player are usually sufficient. We start with the computation of point estimates and finish with the estimation of standard errors.

6.1 Computationally Convenient Point Estimates

Another way to express the recombinant estimator is via iterated expectations. Recall that P represents the set of all hypothetical groupings of the observed strategies. Let P_i be the subset of P containing only those groups which include player i. That is, $\mathrm{P}_i = \{p_j \in \mathrm{P} | i \in p_j\}$. Let J_i represent the number of elements in P_i. Now define $E\{y|i\}$ as the conditional expectation of y restricted to groups containing player i. An estimate of this expectation is:

$$\hat{E}\{y|\ i\} = \left(\frac{1}{\mathrm{J}_i}\right) \sum_{j:p_j \in \mathrm{P}_i} y_j\ . \tag{6-1}$$

Express the recombinant estimator as the average of these estimated conditional expectation over all players i:

$$\hat{\mu} = \left(\frac{1}{m}\right) \sum_{i=1}^{m} \hat{E}\{y|\ i\}. \tag{6-2}$$

As written in (6-2), each group of players is in this summation exactly k times. However, counting every combination k times does not alter the average value.

Consider using $r \leq \mathrm{J}_i$ different combinations involving each player to estimate $E\{y|i\}$. Then, (6-2) becomes

$$\hat{\mu}_r = \left(\frac{1}{m}\right)\sum_{i=1}^{m}\left(\frac{1}{r}\right)\sum_{j:p_j \in R_i} y_j \ . \tag{6-3}$$

where R_i contains r random draws from P_i. When r = 1, $\hat{\mu}_r$ reduces to the baseline estimator, which involves no recombinations at all. Clearly, r = 1 gives a poor estimate of $E\{y|i\}$. Including more recombinations allows the researcher to better estimate $E\{y|i\}$ for each player. However, the gain to recombining is limited by the fact that all it provides is a better estimate of $E\{y|i\}$; it does not increase the number of conditional expectations considered. Therefore, the variance of the estimator is bounded below by the variance of the estimator in which all of the conditional expectations are known:

$$\mu^* = \left(\frac{1}{m}\right)\sum_{i=1}^{m} E\{y|\ i\} \tag{6-4}$$

Thus, once r is large enough such that the variance of $\hat{\mu}_r$ is close to this lower bound, then there is little gain to considering additional recombinations of the data.

To get an idea of how large r needs to be, consider the variance of $\hat{\mu}_r$. The probability that p_j and p_i share at least one player converges to k^2/m from above, as n goes to infinity.[8] So

$$\mathrm{var}(\hat{\mu}) \approx \frac{\sigma^2}{mr} + \frac{k^2}{m}\varphi = \frac{\sigma^2}{krn} + \frac{k\varphi}{n} \tag{6-5}$$

where ≈ denotes approximate equality when the sample size is large. When all possible combinations involving a particular player are used, then r is of order m^{k-1}, resulting in the first term in (6-5) being of lower order. However, if r is fixed, then the first term goes to zero at the same rate as the second, so we have a less efficient estimator. In particular, the asymptotic efficiency of the recombinant estimator with r fixed relative to the full-blown recombinant estimator is $1 + \sigma^2/(rk^2\varphi)$.

[8] The exact probability that two randomly selected games share at least one player is $1 - \prod_{i=0}^{k}(m-k-i)/(m-i)$.

The exact gain from increasing the number of combinations considered is a function of the ratio between φ and σ^2. Table 4 displays the relative efficiencies of the baseline, recombinant and restricted estimators with r fixed at 50, 100 and 500. The table illustrates two points. First, as the covariance across combinations with at least one shared player decreases relative to the variance of the group outcome function, the potential gain from recombining data increases. Second, since the gain to recombining data is great when the covariance is small, there is a greater loss associated with restricting the estimator to consider r recombinations when the covariance is small, as illustrated by the final column of the above table. In particular, consider the case where there are two players in the game and the covariance is 0.01. Here, restricting the estimator to 100 recombinations results in a 25% efficiency loss relative to the fully recombinant estimator. However, we still capture 97.5% of the 98% efficiency gain relative to the baseline estimator.

6.2 Computationally Convenient Standard-Error Estimates

Computation of standard errors for the recombinant estimator requires separately estimating both the variance of the outcome and the average covariance when there is at least one shared player, σ^2 and φ, respectively. Both σ^2 and φ can themselves be estimated via recombinant estimation, because they are themselves means of different group outcomes. To see this result, define a new outcome as the squared deviation from the average value of the old outcome, i.e.

$$s_j = \left(y_j - \hat{\mu}\right)^2 . \tag{6-6}$$

Then the average value of this new outcome is the sample variance of the original outcome variable y, which makes it an estimate of the population variance σ^2. Since we have now posed the estimation of σ^2 as the estimation of a mean value, we can use the recombinant technique to estimate σ^2 more efficiently, computing this mean across all possible k-player groups. Equivalently, one generates a recombinant estimate of σ^2 by generating the same J different hypothetical outcomes used to compute $\hat{\mu}$, and taking their sample variance (whereas $\hat{\mu}$ was their sample mean).

Recombinant estimation can also be used to estimate φ. To see this, let $\Omega = \{p_i, p_j \in \mathrm{P} |\ p_i \cap p_j \neq \varnothing\}$. Define a group of players as $\omega_{ij} = \{p_i, p_j\} \in \Omega$ and the outcome as:

$$h_{ij} = (y_i - \hat{\mu})(y_j - \hat{\mu}). \tag{6-7}$$

Then φ equals the expected value of this new outcome, restricted to observations where at least one player overlaps between groups i and j. Thus, φ is the expected value of a new type of group outcome in which the group size is $2k - 1$ players (two different groups of size k which are constrained to overlap by at least one player). Since φ is the average value of a new type of group outcome, it can be estimated with the recombinant technique. In practice, to estimate φ one generates two columns of data, each row of which contains two different hypothetical outcomes that overlap by at least one player. After generating all such possible pairs of outcomes, one takes the sample covariance between the two columns in order to estimate φ.

In general, recombinant estimation of φ is the most computationally expensive task we consider, because the number of such pairs of outcomes is of order n^{2k-1}, rather than of order n^k as required to estimate μ or σ^2. Thus, in practical computation of standard errors, it is essential to restrict attention to a subset of the possible combinations as was discussed in Section 6.1. In practice, we recommend generating at least 100 random pairs of outcomes overlapping in player 1, 100 pairs overlapping in player 2, and so on up to player m then computing the sample covariance across all such pairs.

Because our estimate of σ^2 comes from recombinant estimation of an outcome (see (6-6)) that is nonlinear in the strategies, this estimate is strictly more efficient (see Proposition 2) than the standard non-recombinant estimate of σ^2 taken across only observed groups. If this result leads to a more accurate estimate of the standard error (recall that we have to estimate φ, as well as σ^2), then we will have even greater power in hypothesis testing. This result, although not fully developed here, is analogous to the power in a *t*-test. Recall that in a *t*-test, increasing the degrees of freedom (while holding the point estimate and the population standard deviation constant) also increases the power of the test. However, as with a *t*-test, the gain in power to increasing the precision of the estimate of the variance is decreasing in the precision of the original estimate. Thus, when the sample size is large, the recombinant and the simple estimators of the variance will produce tests of the same power (just as a *t*-test with 500 degrees of freedom has essentially the same power as one with 1000 degrees of freedom).

7. Concluding Remarks

We have presented a new econometric estimator for use in applications where the quantity of interest is a group outcome based upon the individual strategies of the players. Our recombinant estimator considers all possible hypothetical groupings, yielding considerable efficiency improvements relative to the standard estimator that considers only the observed groupings. In the applied examples we have examined so far, we find efficiency improvements equivalent to increases in the sample size of between 25 and 200 percent. Even greater efficiency improvements are possible in principle, but the exact amount of improvement depends on both the functional form of the group outcome function and on the distribution of the population of individual strategies. We have shown our estimator guarantees some efficiency improvement so long as the group outcome function is not additively separable over the individual strategies.

Experiments on auctions and other normal-form games provide the most straightforward applications of our technique. Recombinant estimation can yield lower-variance estimates of such quantities as auction revenues, market efficiency, or the fraction of the time players achieve the Nash equilibrium in a matrix game. This technique may also help rescue data that might otherwise have become unusable when an experiment goes wrong. For example, suppose that in a twelve-person, simultaneous-move market experiment, one of the twelve participants leaves before submitting a strategy. The market efficiency for that group of twelve subjects cannot be calculated, because one can't find the actual allocation when one of the strategies is missing. However, the other eleven players all submitted valid strategies, so rather than throwing out these observations, one may recombine them with the data from other sessions in order to yield market efficiency estimates.[9]

The estimator also holds some promise beyond simultaneous-move game experiments, and even beyond the field of economics. For example, consider a medical statistician who wishes to estimate the average ratio between healing time with drug A and healing time with drug B. Half the patients in the study received drug A, while the other half received drug B. This situation is isomorphic to the estimation of a two-player game, where the strategies are healing times and the outcome is the ratio between the healing times of one patient and another. The recombinant estimator considers all possible pairings of A-types with B-types, producing an estimate (and a standard error) for this ratio. The technique might be used in any statistical situation that involves matching or grouping of observations.

Future research might be able to generalize our results, for example by exploring the properties of recombinant estimation when the econometrician expects within-group fixed effects in repeated observations of the same group, or by finding additional applications for recombinant estimation. Most importantly, we hope that our work, with its discussions of the practical aspects of computation and of experimental design for use with this technique, will enable other researchers to improve the efficiency of their estimates of behavior in simultaneous-move games. We hope to make this easy for other researchers by providing the software described in footnote 1.

[9] We are grateful to Yan Chen for bringing this application to our attention.

8. Appendix A

This appendix contains a proof of Proposition 1, that the recombinant estimator is weakly asymptotically more efficient than the baseline estimator. Proposition 2, that the recombinant estimator is strictly more efficient when the group outcome function is not additively separable in players' strategies, is proved as a special case. One lemma and two properties of functions of independent random variables are invoked in the proof. Maintaining the notation of section 2, let

$$\varphi \equiv E\{\text{cov}(y_i, y_j) \mid p_i \cap p_j \neq \varnothing \text{ and } p_i \neq p_j\}.$$

The lemma states that the average covariance between groups with at least one common player converges to the average covariance between groups with exactly one common player. More precisely,

Lemma 1: $\lim_{n\to\infty} \varphi = \psi$ where $\psi \equiv E\{\text{cov}(y_i, y_j) \mid \exists \text{ exactly one pair } (t, \tau)\text{: } p_{it} \neq p_{i\tau}\}$.

Proof: The number of pairs $\{y_i, y_j\}$ with exactly r repeated players is of order n^{2k-r}. Therefore, the proportion of the population of pairs with shared players that share exactly one player converges to one as n approaches infinity. *QED*

The two properties of functions of independent random variables are as follows:

1) Let $\mathbf{x}$ be a random k-dimensional vector where x_i is independent of x_j for all $i \neq j$. Then, for all $f(\mathbf{x})$: $R^k \to R$, f can be expressed as

$$f(\mathbf{x}) = f_1(x_1) + \ldots + f_k(x_k) + \delta(\mathbf{x})$$

where x_i is correlated with $f_i(x_i)$ and uncorrelated with all other terms in f. Furthermore,

$$\sum_{i=1}^{k}\left[\text{corr}\left(f(\mathbf{x}),\ f_i(x_i)\right)\right]^2 \leq 1$$

with equality if and only if $\delta(\mathbf{x}) = 0$.

2) Let x be a random variable, $\mathbf{y}$ and $\mathbf{z}$ be k-dimensional random vectors where x, $\mathbf{y}$ and $\mathbf{z}$ are independent of each other. Then, for all $f(\mathbf{x})$: $R^{k+1} \to R$,

$$\left[\text{corr}\left(f(x,\mathbf{y}),\ f(x,\mathbf{z})\right)\right]^2 = \left[\text{corr}\left(f(x,\mathbf{y}),\ f_x(x)\right)\right]^2\left[\text{corr}\left(f(x,\mathbf{z}),\ f_x(x)\right)\right]^2$$

where $f_x(x)$ is defined as in property (1).

Armed with these three results, we are prepared to prove Proposition 1, that the recombinant estimator is weakly more efficient than the standard estimator. In the proof, we point out where the lack of additive separability in the group outcome function leads to a strictly more efficient estimator, proving Proposition 2. Finally, we state Proposition 1 inclusive of all the assumptions in the paper, noting that the relative efficiency of the recombinant estimator is $k\varphi/\sigma^2$.

PROPOSITION 1: Let x_i be independently and identically distributed random variables and $y_j = g\left(x_{p_{j1}}, \ldots, x_{p_{jk}}\right): R^k \to R$. Then, $k\varphi$ is asymptotically less than or equal to σ^2.

Proof: By lemma 1, it is sufficient to prove that the claim holds true when we restrict attention to pairs of outcomes that share exactly one player.

Recall that each player is only repeated in the role in which she is observed and that the game is symmetric within each player role. Therefore, without loss of generality, assume that if a player participates in both y_i and y_j, she is in the same position in both games. There are k player positions, so there are k equally likely types of terms that share exactly one player with y_i in this manner: y_j^t for t = 1 to k, where y_j^t indicates that a player from the t^{th} position has been repeated in game j. (Each type is equally likely because the set of players is restricted to unique combinations where the order of players within a player role does not matter.) Note that y_j^t is independent of y_j^τ for all $t \neq \tau$ since the common player with group i differs and the remaining players are drawn at random.

Let $\rho_{ij,t} = \operatorname{corr}\left(y_i, y_j^t\right)$. By Property (2),

$$\rho_{ij,t}^2 = \rho_{gg_t}^2 \rho_{gg_t}^2 = \rho_{gg_t}^4 .$$

Therefore, the sum of the correlations over the k-player positions is

$$\sum_{t=1}^{k} \rho_{ij,t} = \sum_{t=1}^{k} \sqrt{\rho_{ij,t}^2} = \sum_{t=1}^{k} \rho_{gg_t}^2 \leq 1$$

where the final inequality is a direct application of Property (1).

Note that the final inequality above becomes strict if $y_j = g\left(x_{p_{j1}}, \ldots, x_{p_{jk}}\right)$ cannot be expressed as an additively separable function of the elements of **x**, i.e. $\delta(\mathbf{x}) \neq 0$ in the decomposition in property (1). Furthermore, a strict inequality will carry through to a strict inequality in the statement of Proposition 1. This means a strictly more efficient estimator, and thus proves Proposition 2.

Returning to the proof of Proposition 1, algebraic manipulation of the above inequality yields the desired result, i.e.

$$
\begin{aligned}
& 1 \geq \sum_{t=1}^{k} \rho_{ij,t} = \sum_{t=1}^{k} \frac{\operatorname{cov}\left(y_i, y_j \middle| p_{it} = p_{jt}\right)}{\sqrt{\operatorname{var}(y_i)\operatorname{var}(y_j)}} = \sum_{t=1}^{k} \frac{\operatorname{cov}\left(y_i, y_j \middle| p_{it} = p_{jt}\right)}{\operatorname{var}(y)} \\
& \Rightarrow \sum_{t=1}^{k} \operatorname{cov}\left(y_i, y_j \middle| p_{it} = p_{jt}\right) \leq \operatorname{var}(y) \\
& \Rightarrow \frac{1}{k}\sum_{t=1}^{k} \operatorname{cov}\left(y_i, y_j \middle| p_{it} = p_{jt}\right) \leq \frac{\operatorname{var}(y)}{k} \\
& \Rightarrow \varphi \leq \frac{\operatorname{var}(y)}{k}
\end{aligned}
$$

Since $\lim_{n \to \infty} \varphi = \frac{1}{k}\sum_{t=1}^{k} \operatorname{cov}\left(y_i, y_j \middle| p_{it} = p_{jt}\right)$ by lemma 1 and the fact that each type of covariance in the summation is equally likely.

9. Appendix B

Since the ten auction treatments considered in LLR are all two-player symmetric games, it is sufficient to consider the first treatment, Vickrey auctions of the Barry Sanders card. The data from this experiment are listed in the following table.

Bids from the Barry Sanders 1989 Vickrey Auctions
Each player's bids are listed and players are paired as they were observed in the data.

	First Bid	Second Bid		First Bid	Second Bid		First Bid	Second Bid
Player 1	40	30	Player 13	60	60	Player 25	85	68
Player 2	60	45	Player 14	45	30	Player 26	20	20
Player 3	62	20	Player 15	30	0	Player 27	55	30
Player 4	45	25	Player 16	45	25	Player 28	40	20
Player 5	70	55	Player 17	75	70	Player 29	50	34
Player 6	75	60	Player 18	80	50	Player 30	25	25
Player 7	85	20	Player 19	50	0	Player 31	32	30
Player 8	10	3	Player 20	75	10	Player 32	45	20
Player 9	10	5	Player 21	45	45	Player 33	20	10
Player 10	70	10	Player 22	25	2	Player 34	25	25
Player 11	75	25	Player 23	100	50			
Player 12	55	10	Player 24	78	48			

Furthermore, we restrict our attention to the revenue generated by these auctions, which is the sum of the lowest two bids in each auction.

There were 17 auctions in this format with two bidders in each auction, yielding 34 players. P, the set of all possible combinations of two players, has J = 34(34 – 1)/2 = 561 unique elements. The point estimate is computed by taking the average of the revenue over these 561 auctions, which is equivalent to computing the double summation below.

$$\hat{\mu} = \frac{1}{561}\sum_{i=1}^{J} y_i = \frac{1}{561}\sum_{i=1}^{33}\sum_{j=i+1}^{34} y_{\{i,j\}} = 49.17\,.$$

In the above equation, $y_{\{i,j\}}$ is the revenue generated by the auction involving players i and j.

The 561 outcomes, $\{y_1, \ldots, y_{561}\}$, can be used to estimate both the variance of a single game and the covariance between two games with one shared player. The variance is

$$\hat{\sigma}^2 = \frac{1}{(561-1)}\sum_{i=1}^{J}(y_i - \hat{\mu})^2 = \frac{1}{(561-1)}\left[\sum_{i=1}^{J} y_i^2 - 561\hat{\mu}^2\right] = 736.14$$

where the second representation is computationally more convenient. The covariance is

$$\hat{\varphi} = \frac{\sum_{i=1}^{J-1}\sum_{k=i+1}^{J} 1(p_i \cap p_k \neq \varnothing)\left[(y_i - \hat{\mu})(y_k - \hat{\mu})\right]}{\sum_{i=1}^{J-1}\sum_{k=i+1}^{J} 1(p_i \cap p_k \neq \varnothing)} = 296.56$$

where $1(p_i \cap p_k \neq \varnothing)$ is equal to one if games i and k share a common player and zero otherwise.

Then, $\hat{\sigma}^2$ and $\hat{\varphi}$ are substituted into (2-4) for σ^2 and φ to yield

$$\text{var}(\hat{\mu}) = \frac{2}{m(m-1)}\hat{\sigma}^2 + \frac{4(m-2)}{m(m-1)}\hat{\varphi} = \frac{\hat{\sigma}^2 + 64\hat{\varphi}}{561} = \frac{736.14 + 64(296.56)}{561} = 35.14,$$

which implies that the standard error of the mean is 5.93.

Finally, the recombinant estimate of the variance was used to estimate the variance of the baseline estimator (736.14/17 = 43.3).

For more information, download the Microsoft Excel workbook that implements our estimator (see footnote 1), which provides the data from this application as an example.

References

Friedman, D., 1996. Equilibrium in Evolutionary Games: Some Experimental Results. *The Economic Journal* 106, 1 – 25.

Kagel, J. H., Harstad, R. M., Levin, D., 1987. Information Impact and Allocation Rules in Auctions with Affiliated Private Values: A Laboratory Study. *Econometrica* 55, 1275 – 1304.

List, J., Lucking-Reiley, D., 2000. Demand Reduction in Multi-Unit Auctions: Evidence from a Sportscard Field Experiment. *American Economic Review* 90, 961 – 972.

Mehta, J., Starmer, C., Sugden, R., 1994. The Nature of Salience: An Experimental Investigation of Pure Coordination Games. *American Economic Review* 84, 658 – 673.

Milgrom, P. R., Weber, R. J., 1982. A Theory of Auctions and Competitive Bidding. *Econometrica* 50, 1089 – 1122.

Mitzkewitz, M., Nagel, R., 1993. Experimental Results on Ultimatum Games with Incomplete Information. *International Journal of Game Theory* 22, 171 – 198.

Moir, R., 1998. A Monte Carlo Analysis of the Fisher Randomization Technique: Reviving Randomization for Experimental Economists. *Experimental-Economics* 1, 87 – 100.

Table 1

Proportion of Auctions with the Two Units Split Between the Two Bidders:
A Comparison of the Baseline and Recombinant Estimators

(Standard Errors in Parentheses; P-Values in Double Parentheses)

	Vickrey		Uniform		Vickrey Minus Uniform	
Trading Card	LLR	Recombinant	LLR	Recombinant	LLR	Recombinant
Barry Sanders 1989 Score; BV = $70	0.56	0.53	0.85	0.87	-0.29	-0.34
Nondealers	(0.12)	(0.07)	(0.08)	(0.06)	(0.14)	(0.09)
					((0.034))	((0.000))
Cal Ripken 1982 Topps; BV = $70	0.23	0.30	0.87	0.88	-0.63	-0.58
Dealers	(0.10)	(0.06)	(0.08)	(0.05)	(0.13)	(0.08)
					((0.000))	((0.000))
Jordan 1989 Hoops; BV=$3	0.32	0.34	0.54	0.52	-0.22	-0.17
Nondealers	(0.09)	(0.05)	(0.09)	(0.06)	(0.13)	(0.08)
					((0.079))	((0.036))
Montana 1982 Topps; BV = $3	0.57	0.50	0.45	0.56	0.12	-0.06
Dealers	(0.14)	(0.09)	(0.15)	(0.10)	(0.20)	(0.13)
					((0.565))	((0.675))
Montana 1982 Topps; BV=$3	0.47	0.48	0.53	0.54	-0.07	-0.07
Nondealers	(0.11)	(0.07)	(0.11)	(0.08)	(0.16)	(0.11)
					((0.678))	((0.562))

Table 2
Average Revenue from Vickrey and Uniform Sealed Bid Auctions:
A Comparison of the Baseline and Recombinant Estimators
(Standard Errors in Parentheses; P-Values in Double Parentheses)

	Vickrey		Uniform		Vickrey Minus Uniform	
Trading Card	LLR	Recombinant	LLR	Recombinant	LLR	Recombinant
Barry Sanders 1989 Score; BV = $70	51.06	49.17	48.71	45.25	2.35	3.92
Nondealers	(6.58)	(5.93)	(6.69)	(5.87)	(9.39)	(8.35)
					((0.802))	((0.639))
Cal Ripken 1982 Topps; BV = $70	72.87	73.52	76.13	74.30	-3.27	-0.77
Dealers	(6.56)	(5.98)	(5.42)	(4.72)	(8.51)	(7.62)
					((0.701))	((0.919))
Jordan 1989 Hoops; BV=$3	1.13	1.09	1.71	1.72	-0.58	-0.63
Nondealers	(0.26)	(0.22)	(0.29)	(0.25)	(0.39)	(0.33)
					((0.142))	((0.055))
Montana 1982 Topps; BV = $3	2.37	2.16	2.13	2.14	0.24	0.02
Dealers	(0.32)	(0.29)	(0.44)	(0.36)	(0.54)	(0.46)
					((0.659))	((0.959))
Montana 1982 Topps; BV=$3	0.66	0.65	0.83	0.88	-0.17	-0.23
Nondealers	(0.19)	(0.16)	(0.28)	(0.22)	(0.34)	(0.27)
					((0.613))	((0.403))

Table 3
Mean Expected Payoffs for Demanders and Repliers in the Ultimatum Game
(Standard Errors in Parentheses)

	Cake					
	1	2	3	4	5	6
Player A						
Point Estimate	0.783	1.441	1.874	1.945	1.916	1.816
	(0.014)	(0.031)	(0.043)	(0.049)	(0.059)	(0.065)
Efficiency Gain	0.287	0.383	0.555	0.655	0.626	0.633
Player B						
Point Estimate	0.167	0.327	0.525	0.910	1.248	1.528
	(0.012)	(0.022)	(0.027)	(0.034)	(0.047)	(0.064)
Efficiency Gain	0.164	0.210	0.262	0.388	0.404	0.451

Table 4
Relative Efficiency of the Computationally Restricted Estimator

Number of Players	phi/sigma^2	Recombinant/ Baseline	Restricted/ Baseline	Restricted/ Recombinant
50 Recombinations				
2	0.250	0.500	0.510	1.020
2	0.100	0.200	0.210	1.050
2	0.010	0.020	0.030	1.500
5	0.100	0.500	0.504	1.008
5	0.010	0.050	0.054	1.080
100 Recombinations				
2	0.250	0.500	0.505	1.010
2	0.100	0.200	0.205	1.025
2	0.010	0.020	0.025	1.250
5	0.100	0.500	0.502	1.004
5	0.010	0.050	0.052	1.040
500 Recombinations				
2	0.250	0.500	0.501	1.002
2	0.100	0.200	0.201	1.005
2	0.010	0.020	0.021	1.050
5	0.100	0.500	0.500	1.001
5	0.010	0.050	0.050	1.008