# Enriching Interactions: Incorporating Outcome Data into Static Discrete Games*

Paul B. Ellickson[†]
Duke University

Sanjog Misra[‡]
University of Rochester

November 26, 2007

**Abstract**

When modeling the behavior of firms, marketers and micro-economists routinely confront complex problems of strategic interaction. In competitive environments, firms make strategic decisions that not only depend on the features of the market, but also on their beliefs regarding the reactions of their rivals. Structurally modeling these interactions requires formulating and estimating a discrete game, a task which, until recently, was considered intractable. Fortunately, two-step estimation methods (imported from the auction literature) have cracked the problem, fueling a growing literature in both marketing and economics that tackles a host of issues from the optimal design of ATM networks to the choice of pricing strategy. However, most existing methods have focused on only the discrete choice of actions, ignoring a wealth of information contained in post-choice outcome data and severely limiting the scope for performing informative counterfactuals or identifying the deep structural parameters that drive strategic decisions. The goal of this paper is to provide a method for incorporating post-choice outcome data into static discrete games of incomplete information. In particular, our estimation approach adds a selection correction to the two-step games approach, allowing the researcher to use revenue data, for example, to recover the costs associated with alternative actions. Alternatively, a researcher might use R&D expenses to back out the returns to innovation.

*The authors would like to thank participants at the 2007 Choice Symposium at the University of Pennsylvania and the 2007 QME conference at the University of Chicago. The authors have benefitted from conversations with Pat Bajari, Jeremy Fox, Han Hong, and Harikesh Nair. All remaining errors are our own.

[†]Department of Economics, Duke University, Durham NC 27708. Email: paul.ellickson@duke.edu.

[‡]Corresponding author. William E. Simon School of Business Administration, University of Rochester, Rochester, NY 14627. Email: misra@simon.rochester.edu.

# 1 Introduction

By focusing on the nexus between firms and consumers, marketers and micro-economists continually confront problems of strategic or social interaction. The decision of where to locate a retail store not only depends on the specific capabilities of the firm in question and the customers to whom it wishes to sell its products, but also on the strategic reactions of its rivals. Similarly, the decision of which gym to join may depend on whether a person's spouse or friends belong, as well as how often they intend to go. Not surprisingly, structural models of strategic and social interaction are gaining traction in both fields, providing insight into a host of issues from the optimal design of ATM networks to the decision of whether and with whom to play golf. However, due to both limitations in data availability and constraints inherent in the modeling approach, researchers have almost exclusively focused on discrete outcomes, treating profit or utility as a latent variable, usually parameterized via a reduced form. While this makes efficient use of the often limited data at hand, it can severely limit the usefulness of the model for performing informative counterfactuals or identifying the deep parameters that drive these strategic interactions. For example, in their study of supermarket pricing behavior, Ellickson and Misra (2006) found strong evidence that supermarket chains favor strategies that accord with their rivals, but were unable to pin down exactly *why* such assortative matching was in fact beneficial to the firms. The goal of this paper is to provide a method for incorporating additional, post-choice outcome data into structural models of static discrete games that would allow the researcher to obtain more nuanced insights into a firm's decision making and the nature of strategic interactions across firms.

Our approach combines recent two-step methods for estimating discrete games (Bajari et al. (2006)) with the broader literature on selectivity (Heckman (1979), Heckman and Robb (1985)), as well as recent insights from the literature on moment inequality estimation (Pakes et al. (2006)). The basic idea is quite simple: with data on either revenues or costs, it is possible to recover the strategic parameters that govern profitability. However, to do so we must first correct for the selectivity stemming from the fact that we observe only the revenue or costs (or, more generally, profits) for the action that was actually chosen. We can then use the selectivity corrected estimating equations to construct both revenue and cost counterfactuals. A final inequality estimator can then used to recover the structural parameters that characterize both revenues and costs. In the simplest, fully parametric version of our approach, estimation consists of four steps requiring nothing more complicated that multinomial logits and linear regressions. Since the initial stage of the estimation can be performed either market by market or firm by firm, our methods are robust to multiplicity and can be easily extended to account for heterogeneity, given access to sufficiently rich data.

We illustrate our estimation technique with an empirical exercise that extends on an earlier analysis of supermarket pricing strategies. In a recent paper, Ellickson and Misra (2006) model a supermarket's pricing decision as a discrete, store level game between rival chains. They then estimate the parameters of a reduced form profit function using techniques developed in Bajari et al. (2006). While they find strong evidence of assortative matching by strategy (i.e. firms prefer to offer Every Day Low Prices (EDLP) when they expect their rivals to do likewise), they are unable to pin down exactly why such a strategy is profitable (i.e. is it revenues? or costs? or a combination of both?). However, by including data on revenues (along with the discrete strategic choice of pricing strategy), we are now able to do so. We find that firms coordinate primarily to economize on costs. For example, firms that choose EDLP typically take a revenue hit relative to a PROMO strategy. However, the cost savings associated with EDLP outweigh the expected revenue loss, making it profitable on balance. Further, we also find that coordination of strategies across firms is mostly due to cost savings, as opposed to demand or revenue considerations. Finally, preliminary results suggest that these cost savings are driven by access to common suppliers.

The paper is organized as follows. Section 2 presents a general model of strategic interaction and outlines our four step approach for selectivity-corrected estimation. Section 3 presents our empirical exercise and Section 4 concludes.

## 2 A Model of Strategic Interactions

### 2.1 The Basic Setup: Profit maximization

Although we will introduce a more complex notation for the subsequent empirical exercise, we work here with a simplified set-up that closely follows that of Bajari et al. (2006). We assume that, in each market (whose subscript we suppress for brevity), there are a finite number of players ($i = 1, ..., n$) each choosing a discrete action $a_i \in \{0, 1, ..., K\}$ simultaneously from a finite set. The set of possible action profiles is then $A = \{0, 1, ..., K\}^n$ with generic element $a = (a_1, ..., a_n)$, while the vector of player $i$'s rivals' actions is then $a_{-i} = (a_1, ..., a_{i-1}, a_{i+1}, ..., a_n)$.

The state vector for player $i$ is denoted $s_i \in S_i$, while the state vector of all $n$ players is $\mathbf{s} = (s_1, ..., s_n) \in \Pi_i S_i$. The state vector $\mathbf{s}$ is known to all firms and observed by the econometrician. It describes features of the market and characteristics of the firms that we assume are determined exogenously. For each firm, there are also two privately observed state variables. While each firm perfectly observes its own private state variables, they are known to the econometrician and rival firms only in distribution. These asymmetrically observed state variables are denoted $\epsilon_i^R(a_i)$ and $\epsilon_i^C(a_i)$, or more compactly $\epsilon_i^R$ and $\epsilon_i^C$, and represent firm specific shocks to the revenue ($R$) and cost ($C$) associated with each strategy. We will sometimes write the two element vector of private shocks as $\epsilon_i$. The private information assumption makes this a game of incomplete or asymmetric information and the appropriate equilibrium concept one of Bayes-Nash Equilibrium (BNE). For any given market, the $\epsilon$'s are assumed to be *i.i.d.* across firms and actions. They are drawn from a joint distribution $f\left(\epsilon_i^R, \epsilon_i^C\right)$ that is known to everyone, including the econometrician.

Firms choose strategies with the objective of maximizing expected profits

$$\pi_i\left(\mathbf{s}, a_i = k, a_{-i}, \epsilon_i(k), \eta_i(k); \theta^k\right) = R_i^k - C_i^k, \tag{1}$$

broken out here as revenue minus costs. The revenue and cost terms can be expressed more explicitly as

$$\begin{aligned} R_i^k &= R\left(\mathbf{s}, a_i = k, a_{-i}, \epsilon_i^R(k), \eta_i^R(k); \theta_R^k\right) & (2) \\ C_i^k &= C\left(\mathbf{s}, a_i = k, a_{-i}, \epsilon_i^C(k), \eta_i^C(k); \theta_C^k\right) & (3) \end{aligned}$$

Note that both the revenue and cost terms include measurement errors $\eta_i^R$ and $\eta_i^C$ (or, more compactly, $\eta$), each of which are assumed to be mean zero. Consistent with the information structure of the game, we assume firms choose the action $a_i$ that yields the highest expected profit relative to all alternatives, so that

$$\mathcal{E}\left[\pi_i\left(\mathbf{s}, a_i = k, a_{-i}, \epsilon_i\left(k\right), \eta_i\left(k\right); \theta^k\right) - \pi_i\left(\mathbf{s}, a_i = k', a_{-i}, \epsilon_i\left(k'\right), \eta_i\left(k'\right); \theta^{k\prime}\right)\right] \geq 0 \quad (4)$$

for all $k'$. Note that the expectation is over the actions of rival firms, as well as both measurement errors.

## 2.2 Constructing Firm Beliefs

Because the actions of a firm's rivals enter its payoff function directly, the structure of this discrete game differs from a standard discrete choice model. The game structure transforms the problem into a system of simultaneous, multinomial discrete choice equations. Since this is a game of incomplete information, firms cannot perfectly predict the actions of their competitors. Moreover, since the $\epsilon$'s are treated as private information, a particular firm's decision rule $a_i = \mathcal{E}_\eta\left[d_i\left(\mathbf{s}, \epsilon_i^R, \epsilon_i^C, \eta_i\right)\right]$ is a function of the state vector ($\mathbf{s}$) and its *own* private information, but *not* the private information of its rivals. In other words, each firm's beliefs over what their competitors will do are functions of observables alone. Thus, from the perspective of its rivals, the probability that a given firm chooses action $k$ conditional on the common state vector $s$ is given by

$$P_i\left(a_i = k|\mathbf{s}\right) = \int\int\int \mathbf{1}\left\{d_i\left(\mathbf{s}, \epsilon_i^R, \epsilon_i^C, \eta_i\right) = k\right\} \times f\left(\epsilon_i^R, \epsilon_i^C\right) \times g\left(\eta_i\right) d\epsilon_i^R d\epsilon_i^C d\eta_i \quad (5)$$

where $\mathbf{1}\left\{d_i\left(s, \epsilon_i^R, \epsilon_i^C, \eta_i\right) = k\right\}$ is an indicator function equal to 1 if firm $i$ chooses action $k$ and 0 otherwise. We let $\mathbf{P}$ denote the set of these probabilities. Since the firm does not observe the actions of its competitors prior to choosing its own action, its decisions are based on these expectations. For tractability, we assume the $\epsilon$'s are additively separable and independent across firms and actions. These and other assumptions will be made clear in the sequel.

### 2.3 Incorporating Information on Post Choice Outcomes

To proceed, we must specify the additional information that is available to the researcher beyond simply the discrete choice of strategy. Since revenue data is often easier to obtain than information on costs, we assume that we observe revenues (but not costs). This matches the empirical example that is presented later in the paper. However, the case where costs (or other post choice outcome data) are observed instead is completely analogous.

While the methods presented here are fairly general, we should note that the inclusion of more detailed data (e.g. prices and quantities) would require the researcher to impose some added structure (i.e. specify a demand function). Given this additional structure, the basic estimation approach presented below could be applied. Of course, having actual profit data is also easily accommodated in our framework.

#### 2.3.1 Calibrating Revenue

Given revenue data, we assume that revenue can be decomposed as some function of the state vector and the actions, along with the additive (private information) revenue shock $\left(\epsilon_i^R(k)\right)$ and the additive measurement error $\left(\eta_i^R(k)\right)$. In particular,

$$R_i^k \approx R\left(\mathbf{s}, a_i = k, a_{-i}; \theta_R^k\right) + \epsilon_i^R(k) + \eta_i^R(k)\,. \tag{6}$$

Since the error components ($\eta$ and $\epsilon$) are not empirically distinguishable to the econometrician, the revenue function can equivalently be written as

$$R_i^k \approx R\left(\mathbf{s}, a_i = k, a_{-i}; \theta_R^k\right) + \omega_i^R(k) \tag{7}$$

where $\omega_i^R(k) = \epsilon_i^R(k) + \eta_i^R(k)$. Because we only observe revenues for the strategies that are actually chosen, there is a selectivity problem: since firms choose the strategy that maximizes profits, strategies are not randomly assigned. Consequently, $\mathcal{E}\left(\omega_i^R | a_i = k\right) \neq 0$, although the selectivity bias is clearly driven by $\epsilon_i^R(k)$. Since revenue depends on the choice of strategy, the revenue equation requires a selection correction. We will specify how this is done below. However, before doing so, we must also parameterize the cost function.

#### 2.3.2 Cost Parameterization

We adopt a similar specification for the cost function as for revenue,

$$C_i^k \approx C\left(s, a_i = k, a_{-i}; \theta_C^k\right) + \epsilon_i^C(k) + \eta_i^C(k). \tag{8}$$

Once again, $\epsilon$ is treated as private information while $\eta$ is measurement error. Since costs are unobserved, and due to the "outside good" normalization, we must work with cost differences rather than levels. We adopt the following formulation for these differences

$$\mathcal{E}\left[C_i^k - C_i^{k'}\right] = \Delta C_i\left(k, k'|s, \mathbf{P}, \theta_C\right) + \Delta\epsilon_i^C\left(k, k'\right) \tag{9}$$

where $\Delta\epsilon_i^C(k, k') = \epsilon_i^C(k) - \epsilon_i^C(k')$. Note that since the firm makes strategy decisions based on *expected* revenue and costs, the measurement errors ($\eta$) drop out, allowing actions to be replaced by expectations (probabilities). At the time this decision is made, the measurement errors are unknown to the firm and can be integrated out. Having parameterized the full model, we are now ready to outline our estimation algorithm.

#### 2.3.3 Constructing the Likelihood

Let $\mathbf{\Psi}_i^k$ be an indicator function indicating that firm $i$ chooses action $k$ conditional on all parameter ($\theta$), the vector of state variables ($\mathbf{s}$), its own private information components ($\epsilon$), and the expectations it has over the actions of its competitors $(\mathbf{P})$. This can be written as

$$\mathbf{\Psi}_i^k\left(a_i = k|\theta, \mathbf{P}, \mathbf{s}, \epsilon_i^R, \epsilon_i^C\right) = \sum_{k' \neq k} \mathbf{1}\left(\mathcal{E}_\eta\left[\begin{array}{c} \pi_i\left(\mathbf{s}, a_i = k, \mathbf{P}_{-i}, \epsilon_i(k), \eta_i(k); \theta^k\right) \\ -\pi_i\left(\mathbf{s}, a_i = k', \mathbf{P}_{-i}, \epsilon_i(k'), \eta_i(k'); \theta^{k\prime}\right) \end{array}\right] > 0\right). \tag{10}$$

Let $g\left(\mathbf{R}|\theta_R, \mathbf{s}, \mathbf{a}, \epsilon^R\right)$ be the joint conditional density of revenues generated by the measurement error terms ($\eta$). The conditioning reflects the fact that the observed vector of revenues (across players), $\mathbf{R}$, is conditioned on the state variables ($\mathbf{s}$), revenue function parameters ($\theta_R$), revenue private information components $\left(\boldsymbol{\epsilon}^R\right)$, and the actions[1] of all players ($\mathbf{a}$). Finally, let $k^*$denote observed actions.

Given these pieces, the overall sample likelihood can then be written as

[1] Note that revenues are a direct function of *actual* actions not *expected* actions of competition. The *expected actions* of a firm's competitors impact revenues only via that firm's choice of action.

$$\mathbb{L}=\int\int\left\{g\left(\mathbf{R}|\theta_R,\mathbf{s},\mathbf{a}=\mathbf{k}^*,\epsilon^R\right)\times\prod_i\mathbf{\Psi}_i^{k^*}\left(a_i=k^*|\theta,\mathbf{P},\mathbf{s},\epsilon_i^R,\epsilon_i^C\right)\right\}f\left(\epsilon^R,\epsilon^C\right)d\epsilon^R d\epsilon^C \tag{11}$$

$$\textbf{s.t. } P_i\left(a_i=k\right)=\int\int\mathbf{\Psi}_i^k\left(a_i=k|\theta,\mathbf{P},\mathbf{s},\epsilon_i^R,\epsilon_i^C\right)\times f\left(\epsilon_i^R,\epsilon_i^C\right)d\epsilon_i^R d\epsilon_i^C$$

A few comments on the structure of the likelihood function are warranted. First, note that $\epsilon^R$ and $\theta_R$ enter both $g$ and $\mathbf{\Psi}$. This is similar to the discrete-continuous demand model developed by Hanemann (1984). However, unlike Hanemann's demand framework, our setting involves multiple agents playing a discrete game, with the observed choice probabilities representing the equilibrium outcome of this game. Second, the equilibrium mapping constraint requires that the expectations ($\mathbf{P}$) be equal to equilibrium probabilities $\left(\mathbf{\Psi}_i^k\right)$. This is a fixed point problem that must be evaluated for any guess of the parameter vector $\theta$. This complicates the evaluation of the likelihood since using such a nested fixed point approach (Rust, 1987) is cumbersome and fraught with complications such as the enumeration and selection of fixed points. Finally, we note that even with an equilibrium selection rule in place, maximizing the constrained likelihood is a numerically challenging task. In what follows, we present a sequential estimation algorithm that greatly simplifies the estimation of the parameters.

### 2.4 Methodology: Sequential Estimation

In this section, we describe the general methodology used to obtain estimates of the cost parameters. The intuition behind the algorithm is straightforward: profits are the linear difference between revenues and costs. If we can ascertain the impact of the state variables on revenues and use these revenue equations to construct counterfactual revenues, the conditions for profit maximization allow us to recover the effect of the state variables on costs. Here are the four steps that constitute our estimation algorithm (details follow):

**Step 1:** *Flexibly estimate* $\mathbf{P}$ *as a function of all state variables, yielding* $\widehat{\mathbf{P}}$

Note that it is crucial that the first step provide consistent estimates of the components of $\mathbf{P}$, since these will in turn be used to construct estimates of players' expectations over rivals' actions. This estimate $\left(\widehat{\mathbf{P}}\right)$ is also the key component in the selectivity correction that follows. There are various approaches to constructing $\widehat{\mathbf{P}}$ in a flexible manner. These might include neural networks, the method of sieves, kernel estimators, or series estimators. Ideally, the researcher should use nonparametric methods to implement this step. Of course, if data limitations exist, one may have to employ less data intensive semiparametric or even parametric methods.

**Step 2:** *Estimate selectivity corrected Conditional Revenue Equations $R_i$, obtaining $\widehat{\theta}_R$*

The second step recovers the selectivity corrected structural parameters of the revenue function. The approach we recommend here follows methods proposed by Heckman and Robb (1985) and Ahn and Powell (1993). Implementation is straightforward. We show in the appendix that our approach amounts to running separate revenue regressions for each strategy $(k)$ with an additional component $\Lambda_k\left(\widehat{\mathbf{P}}_i\right)$ which is a flexible function of $\widehat{\mathbf{P}}_i$. Specifically, the regressions are given by

$$R_i^k\left(\mathbf{s}, a_i = k, a_{-i}; \widehat{\theta}_R^k\right) = R\left(s, a_i = k, a_{-i}; \theta_R^k\right) + \tilde{\omega}_i^R(k) + \Lambda_k\left(\widehat{\mathbf{P}}_i\right) \qquad (12)$$

In practice, the control function $\Lambda_k\left(\widehat{\mathbf{P}}_i\right)$ can be approximated using splines or series expansions (polynomials). The main advantage of our approach is that it relies only on $\widehat{\mathbf{P}}_i$ in the correction term and does not require any parametric assumption on the error structure. While this flexibility does come at some cost, the alternative would be to impose particular distributions on the $\epsilon$ and $\eta$ such that the selectivity correction is empirically tractable. This is not a trivial task since (i) we are analyzing a multinomial choice problem and (ii) the same errors appear in both the selection equation (choice model) and in the outcome equation (revenue regression). These issues also distinguish our framework from the standard selectivity approach in which the errors in the two equations are simply assumed to be correlated (See e.g. Mazzeo 2002).[2] We should note here that our problem simplifies greatly due to the private information assumptions made in the set-up of the discrete game. Since the private information components are i.i.d. across players and actions, the fact that these

[2] Note that this does not preclude the econometrician from using specific distributions (say extreme value errors for $\epsilon$) since these assumptions can be imposed in the final step of our algorithm. Being agnostic about these errors at this point simply retains flexibility while also being fairly simple to implement.

actions enter the regression equation does not raise endogeneity issues. In particular, the private information assumption (on $\varepsilon$) allows the joint selectivity problem (the revenues of *all* players are conditioned on the actions of *all* players) to be decomposed into a collection of individual selectivity problems. Of course, the revenues may ex-post be correlated across players on account of the measurement errors ($\eta$'s). In cases where the cardinality of an individual firm's action space is large, the researcher will likely face a dimensionality[3] issue in modeling $\Lambda_k\left(\widehat{\mathbf{P}}_i\right)$. One possible solution is to assume that Dahl's (2002) *index sufficiency* assumption holds, and rely only on $\Lambda_k\left(\widehat{P^k_{iq}} : q \subset K\right)$. This reduction in dimension is, unfortunately, somewhat ad-hoc since it is not based on any utility theoretic primitive.

**Step 3:** *Use $\widehat{\theta}_R$ to construct Counterfactual Revenue Differences $\Delta R_i\left(\widehat{k, k'|s}, \widehat{\mathbf{P}}, \widehat{\theta}_R\right)$ where,*

$$\Delta R_i\left(\widehat{k, k'|s}, \widehat{\mathbf{P}}, \widehat{\theta}_R\right) = R^k_i\left(\widehat{\mathbf{s}, \widehat{\mathbf{P}}; \widehat{\theta}^k_R}\right) - R^{k'}_i\left(\widehat{\mathbf{s}, \widehat{\mathbf{P}}; \widehat{\theta}^{k'}_R}\right).$$

Once the revenue parameters are available, constructing the counterfactual revenue differences (CRDs) is straightforward. The CRDs are simply the fitted revenue differences between the strategy that was chosen ($k$) and those that were not ($k'$). Note here that $\Delta R_i\left(\widehat{k, k'|s}, \widehat{\mathbf{P}}, \widehat{\theta}_R\right)$ is not the complete ex-ante expected revenue difference since it does not include the private information terms. However, this redounds to our advantage since we can now plug these CRD's into the structural choice problem to estimate the cost differences without worrying about the private information issues.[4]

**Step 4:** *Estimate cost function parameters ($\theta_C$) from the profit maximizing constraints.*

Once the CRDs are known, the profit maximizing constraints in (4) can be used to estimate the implied cost differences. The profit maximizing constraints outlined in (4) are equivalent to the following empirical condition:

$$\Delta R_i\left(\widehat{k, k'|s}, \widehat{\mathbf{P}}, \widehat{\theta}_R\right) - \Delta C_i\left(k, k'|s, \widehat{\mathbf{P}}, \theta_C\right) \geq \kappa_i$$

[3] A second order polynomial approximation with $J$ alternatives requires the estimation of $J + \Sigma^J_{i=1} i$ terms. For example, with 5 alternatives one would have to estimate 20 parameters for the selectivity correction component alone.

[4] In other words, simply plugging in revenue *data* for the observed choice is not an option since such data includes realizations of the private information and measurement error shocks. The latter is clearly not in the firm's information set when the discrete choice is being made.

where $\kappa_i = \Delta\epsilon_i^C - \Delta\epsilon_i^R$ represents the difference in the private information components for strategies $k$ and $k'$. Any set of parameters $\widehat{\theta}_C$ that satisfies the above condition yields a consistent estimate of $\theta_C$. The estimation can be performed parametrically (using a multinomial logit, for example) by imposing specific assumptions on $\kappa$ (via appropriate assumptions on the $\epsilon$) or semi-parametrically via maximum score. In principal, a bounds estimator similar to the moment inequalities approach proposed by Pakes et al. (2006) could also be used. The estimation approach outlined here has been kept deliberately general. The actual implementation of the approach would require specific choices on the part of the researcher. In the next section we describe an example that illustrates the application of the described methodology.

## 2.5 Discussion

### 2.5.1 Identification

In empirical static games, as in standard discrete choice models, identification of the latent profit or payoff parameters comes from the covariation between the explanatory variables and the revealed choice data. The identification of the strategic effects, however, is slightly more involved, requiring explicit exclusion restrictions. These exclusion restrictions usually take the form of continuous covariates that impact the payoffs of each individual player directly, but do not influence the payoffs of the other players (except through the expected actions). Given such exclusion restrictions, the strategic effects are identified since the variation in the beliefs a firm has over it's competitors actions are now driven at least partly by covariates that do not shift its own profits directly.[5] Our methodology also requires such exclusion restrictions.

In addition to having exclusion restrictions *across* firms we also need another set of exclusion restrictions *within* each firm. These restrictions are required to identify the revenue regression parameters. Since the selectivity correction term is also a function of state variables, its inclusion into the second step regression poses identification problems. The solution in the selectivity literature (see e.g. Heckman and Honore, 1990) is to have a set of variables that impact the discrete choice but are excluded in the regression specification.

[5] We note here that there is no collinearity problem *per se* since the beliefs are typically nonlinear transformations of payoffs. However, in the absence of exclusion restrictions the identification of strategic effects is based purely on parametric assumptions on the error structure and functional restrictions on the payoffs.

The economic structure of our problem helps in this regard since costs are, by definition, excluded from revenues. More precisely, all we need are variables that are assumed to affect costs but not revenues. Ideal exclusions, in our opinion, are variables that influence fixed costs since these can be assumed in most typical economic models to be independent of demand side constructs. Note that the two sets of exclusion restrictions (within and across firms) could overlap as long as they have some elements that are not common to both.

#### 2.5.2 Inference

Since the multi-step structure of our estimation routine makes standard inference approaches inapplicable, we construct standard errors via the bootstrap. The construction of such a bootstrap procedure requires some care, since simply bootstrapping over observations would result in severe biases (i.e. we might drop a firm from a given market). In our application we bootstrap over markets which is effective since we have many markets that are assumed to be independent of each other. Future research might investigate other approaches such as subsampling or jack-knife methods.

Finally, the consistency of our estimates should follow from the arguments laid out in the selectivity (Ahn and Powell, 1993) and static games (Bajari et al. 2006) literature. While a proof of consistency is beyond the scope of this paper, since our algorithm is simply an extension of standard "two-step" methods, the usual arguments based on Newey and McFadden (1994) apply here as well. In general, as long as the first stage is consistently estimated, each subsequent stage should result in consistent estimates.

## 3 Application: Supermarket Pricing Strategies

To illustrate how our approach can be applied in practice, we extend the empirical model of supermarket pricing strategies introduced in Ellickson and Misra (2006) to incorporate store-level data on revenue. For the sake of brevity, we will provide only a cursory overview of the choice model and dataset here, referring readers to the previous paper for a more detailed description. Pricing decisions are modeled as a static, discrete game of incomplete information in which supermarket firms choose among three pricing strategies: Every Day Low Pricing (EDLP), Promotional Pricing (PROMO), and mixture of the two, commonly known as hybrid pricing (HYBRID). Firms choose the strategy that maximizes expected store-level profits given their beliefs regarding the actions of their rivals. Firms condition their choices on an underlying state variable that includes both store and firm level covariates, as well as market level demographics. Using a two-step estimation procedure based on Bajari et al. (2006) and Aguirregabiria and Mira (2006), Ellickson and Misra (2006) find that stores coordinate on the choice of strategy. For example, stores that choose EDLP expect to earn higher profits when their rivals do likewise. The purpose of the current exercise is to discover why.

### 3.1 Implementation

Our implementation follows the four step procedure described above. In the first step, we recover consistent estimates of the choice probabilities that will be used to construct the control functions employed in step two. Although firms are choosing among just three pricing strategies, the inclusion of a large number of continuous covariates precludes the use of fully flexible non-parametric methods for predicting probabilities (e.g. kernel, sieve, or series estimators). Therefore, we proceed semi-parametrically, estimating the first stage market by market using a flexible multinomial logit specification that includes higher order terms for each covariate, along with a set of bivariate interactions.

The consistent estimates of the choice probabilities (i.e. $\widehat{\mathbf{P}}$) obtained in this first stage are used to construct the control functions that correct for selectivity in the second stage revenue regressions. Since we only model three choices, the curse of dimensionality that can arise in constructing such control functions is not a concern. Therefore, we do not need to make any additional assumptions (e.g. index sufficiency) to reduce the dimensionality. Given the large amounts of data at our disposal, we were able to employ third order polynomials to approximate $\Lambda_k(\mathbf{P})$. We also experimented with higher order terms but found only very small differences in the resulting estimates.

The third step is straightforward, simply requiring construction of predicted and counterfactual revenues, based on the observed covariates and the results of step 2. With these counterfactual revenue differences in hand, we move on to the final step, in which we recover the cost differences that rationalize the observed choices given the expected revenue differences that were constructed in step 3. Here we again follow a semi-parametric approach, this time based on pairwise comparisons between the selected choice and each of the unchosen alternatives. In particular, we use a smoothed pairwise maximum score procedure similar to the methods developed in Fox (2007). The results of an alternative specification based on a simple multinomial logit framework (not reported) were broadly similar. All standard errors were constructed by bootstrapping over markets.

## 3.2 Results and Discussion

### 3.2.1 Step 1: Consistent Estimates of P

As noted above, our first step was carried out using a flexible multinomial logit specification. Since the coefficient estimates from this procedure are not particularly enlightening (and only the fitted values will be used in what follows), we simply note here that 64.8% of stores's pricing strategy were correctly classified by this first stage. This measure of fit is represented visually in Figure 1, which plots the predicted profit differences (i.e. inverted probabilities) for each observation (store) in the dataset. The vertical axis shows HYBRID profits (relative to PROMO) while the horizontal axis plots EDLP (again versus PROMO). The observations are color coded by strategy (HYBRID is orange, EDLP is blue, and HYBRID is black), yielding a clear visual indication of relative fit. This should provide some level of confidence in the selection corrections that follow.

### 3.2.2 Step 2: Revenue Function Estimates

In step 2, we project observed store level revenues onto several covariates characterizing the relative size and attractiveness of a given market, along with two measures of rival actions (the share of rival stores choosing EDLP and PROMO, respectively). These revenue regressions are corrected for selectivity using the control functions described above. The results of this exercise are presented in Table 1. The first three columns contain revenue regressions for EDLP, HYBRID, and PROMO respectively, while the last two contain the coefficients for EDLP and HYBRID differenced against PROMO. With the exception of the number of families[6] and two additional covariates, all regressors are significant at the 1% level. As one might expect, store size, median rent, and median income all have a large,

[6] The market definition used here, which is essentially a ZipCode, does not provide enough variation in population to identify this effect.

positive and significant impact on revenue, irrespective of the choice of pricing strategy. Similarly, both the percentage of minority residents and the median number of vehicles have a negative impact. The latter covariate most likely reflects consumers' ability to search. Note that, consistent with the predictions of Lal and Rao (1997), the PROMO strategy is hurt the least by consumer search, reflecting their view that this "hi-lo" pricing strategy is explicitly aimed at "cherry pickers". Focusing on the strategic effects, we find that, consistent with standard arguments regarding "business stealing" (congestion) effects, the revenue of every strategy is decreasing in the proportion of rivals who choose that same strategy. While this might at first appear to contradict the findings presented in Ellickson and Misra (2006), an examination of the relative effects reveals that this is not in fact the case. For example, although the presence of competing EDLP stores reduces the revenues that a particular store can expect to earn by choosing EDLP, the damage to a store that selects PROMO instead is even greater. However, as we will see later, this revenue differential is small compared to the relative cost advantage of choosing EDLP versus PROMO.

#### 3.2.3 Step 3: Constructing Counterfactual Revenue Differences

Having obtained revenue estimates in step 2, this next step simply involves constructing counterfactual revenue differences - the expected gain (or loss) from choosing an alternative pricing strategy. The results of this exercise are presented in Figure 2, which display histograms for EDLP and HYBRID (versus PROMO) for the stores that actually chose each of these strategies. We see that, amongst stores choosing EDLP, about a third would sustain a revenue hit by switching to PROMO. However, the majority of such stores would actually stand to *gain* revenue by switching to PROMO, implying that substantial cost saving must be working to offset the forgone revenue (we verify this claim in the following section). Interestingly, the analogous data for HYBRID stores are quite different: the vast majority of HYBRID stores would *lose* revenue by switching to PROMO, indicating a more limited role for costs. However, to recover the actual cost implications of choosing HYBRID (and verify our claim regarding EDLP), we must turn to the cost differences recovered in step 4.

### 3.2.4 Step 4: Cost Function Estimates

The final step of our four step procedure involves comparing the counterfactual revenue differences constructed in step 3 to the actual choices that each store made in order to back out the cost differentials that rationalize their observed actions. Before discussing the coefficient estimates from our smoothed pairwise maximum score procedure, we present two figures that illustrate the overall fit of the model along with the primary factors driving each strategic choice. Figure 3 plots both the revenue and cost differences of EDLP versus PROMO for stores that actually chose EDLP. The red line represents the locus of indifference points where the revenue and cost differences coincide. The region with the highest density of observations is the lower left quadrant, to the right of the red line. This means that, relative to PROMO, stores choose EDLP primarily to economize on costs. In particular, for the typical store choosing EDLP over PROMO, the amount they expected to gain in costs outweighed the amount they expect to lose in revenues. In contrast, Figure 4 shows that the decision to choose HYBRID (rather than PROMO) is driven by a combination of *both* revenues and costs. The highest density region is in the lower right quadrant, implying dominance along both dimensions.

Turning next to the coefficient estimates, we can identify the covariates that drive these cost differences. Note that the estimates presented in Table 2 are directly comparable to those presented in the last two columns of Table 1, as everything is now expressed in differences. Firm characteristics enter as expected, with increased store size, chain size, and vertical integration all reducing the costs of both EDLP and HYBRID relative to PROMO. Not surprisingly, the impact of vertical integration is a particularly strong factor in the EDLP/PROMO comparison, reflecting the firm level investments needed to profitably implement EDLP. However, the most interesting results are those involving strategic interaction. The large negative coefficient on the share of EDLP rivals in the first column implies that coordinating on EDLP actually reduces costs relative to PROMO. The same is true for HYBRID. Moreover, the magnitudes of these effects are significantly larger in magnitude than the ones we found on the revenue side, indicating that the primary reason firms are coordinating on pricing strategies is to reduce costs.

The relative size of these effects can be gauged by examining Table 3, which reports average cost, revenue and profit differences for stores that choose EDLP or HYBRID versus PROMO. In particular, choosing EDLP yields an average cost savings of $823,392 over PROMO, which is about 5.75% of average (store-level) revenues. However, this choice yields an average decrease in revenue of $447,356. The net profit difference is $376,063, reflecting the greater magnitude of the cost savings. In contrast, choosing HYBRID yields substantial gains along both dimensions.

# 4 Concluding remarks

This paper proposes a novel approach to incorporating outcome information into static games of incomplete information. A simple four step estimation recipe is provided along with an application outlining implementation details and new insights obtained.

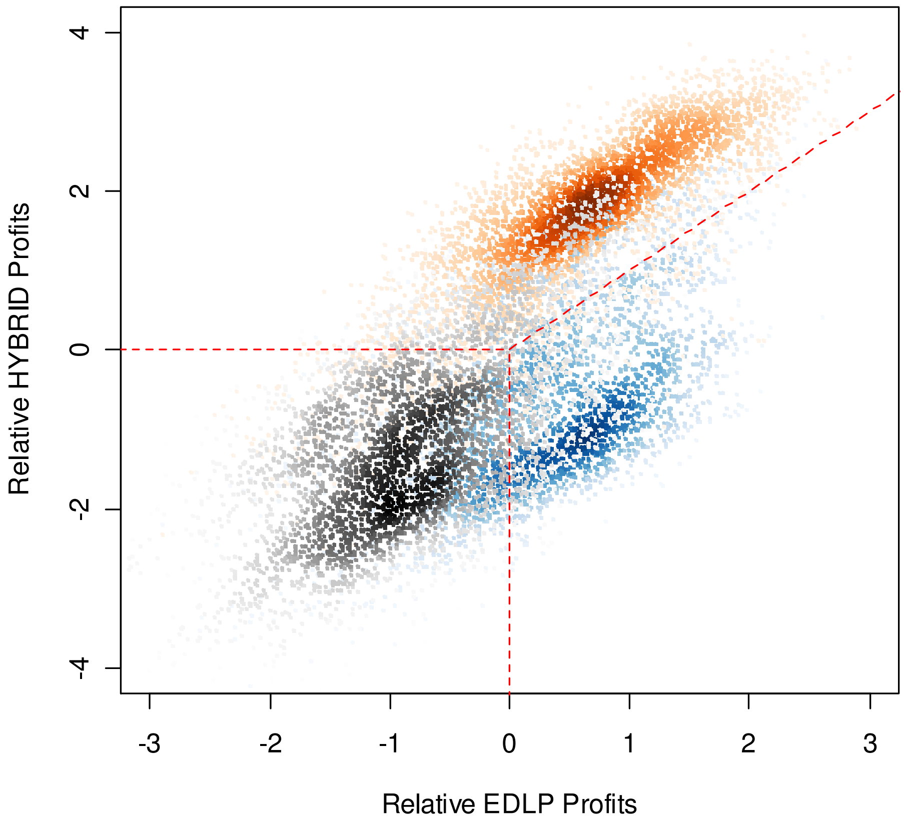


Figure 1: Overall Fit

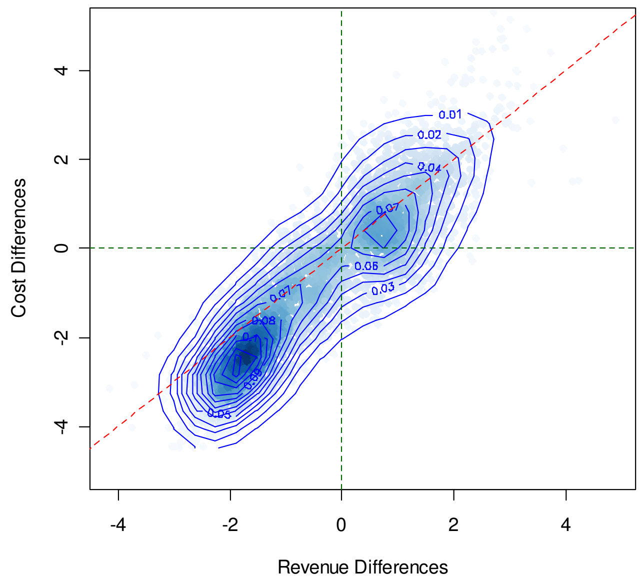


Figure 3: EDLP versus PROMO

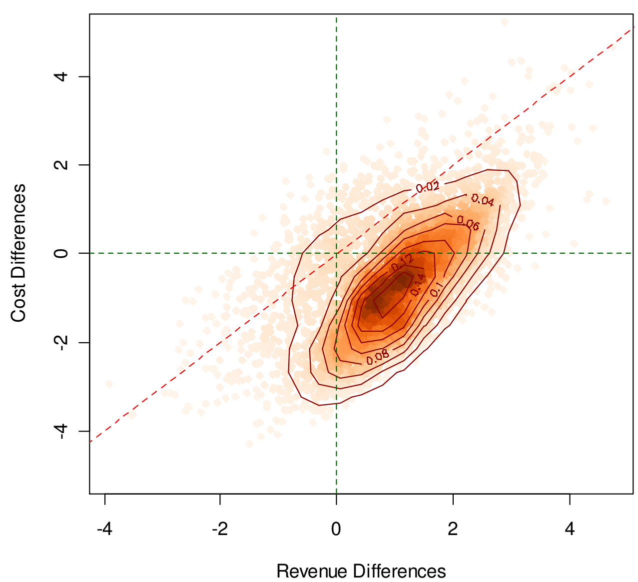


Figure 4: HYBRID versus PROMO

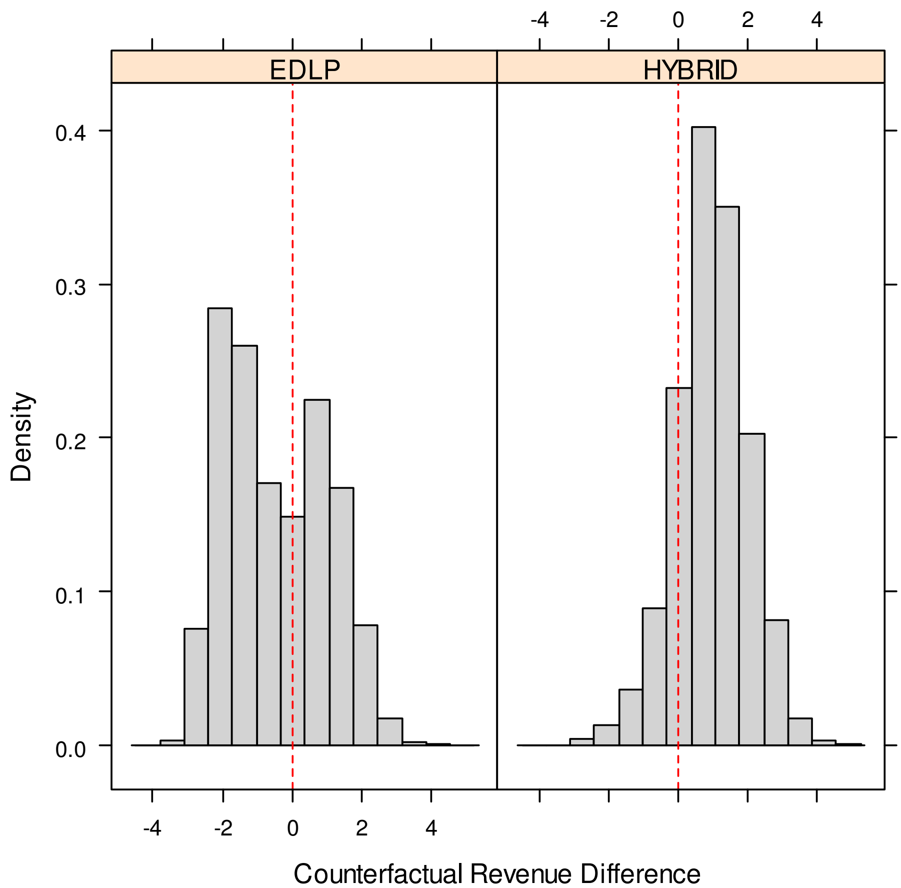


Figure 2: Counterfactual Revenue Differences

Table 1: Selectivity Corrected Revenue Regressions

| **Variable** | **EDLP** | **HYBRID** | **PROMO** | $\beta_E - \beta_P$ | $\beta_H - \beta_P$ |
|---|---|---|---|---|---|
| Intercept | 0.13590 | 0.13576 | 0.05318 | 0.08272 | 0.08259 |
| $\rho^E_{-i_j^{lm}}$ | -0.03493 | -0.09209 | -0.05197 | 0.01703 | -0.04012 |
| $\rho^P_{-i_j^{lm}}$ | 0.00318$^{ns}$ | 0.03482 | -0.06304 | 0.06623 | 0.09786 |
| Number of Families | -0.00016$^{ns}$ | 0.00060$^{ns}$ | -0.00055$^{ns}$ | 0.00039 | 0.00115 |
| Store Size | 0.00756 | 0.00728 | 0.00693 | 0.00062 | 0.00035 |
| Chain Size | -0.00004 | -0.00139$^{ns}$ | 0.00001 | -0.00005 | -0.00140 |
| Proportion Black | -0.04784 | -0.05200 | -0.02420 | -0.02364 | -0.02780 |
| Median Rent | 0.00007 | 0.09557 | 0.00012 | -0.00005 | 0.09545 |
| Median Income | 0.00090 | 0.00145 | 0.00082 | 0.00008 | 0.00063 |
| Median # of Vehicles | -0.05738 | -0.07906 | -0.03010 | -0.02728 | -0.04896 |

All coefficients are significant at the 1% level, with the exception of those denoted “ns”.

Table 2: Cost Estimates

| *Variable* | $\Delta C_{\text{E-P}}$ | $\Delta C_{\text{H-P}}$ |
|---|---|---|
| Intercept | 0.929500 | -2.32910 |
| $\widehat{\rho}_{-i}^{E}$ | -2.225800 | 0.67050 |
| $\widehat{\rho}_{-i}^{P}$ | 3.159600 | 6.47780 |
| Store Size | -0.003423 | -0.00922 |
| Chain Size | $-0.000101^{ns}$ | $-0.00006^{ns}$ |
| Vertically Integrated | -0.862900 | -0.06620 |
| FT Employees | $0.000412^{ns}$ | -0.00092 |

Table 3: Comparison of Strategies

| **Metric (Mean)** | **EDLP** | **HYBRID** |
|---|---|---|
| Cost Savings (%) | 5.75% | 4.20% |
| Cost Savings ($) | 823392 | 672100 |
| Revenues ($MM) | 14.31 | 16.02 |
| Revenue Difference ($) | (447,356) | 953,447 |
| Revenue Difference (% of Revenue) | -3.13% | 5.95% |
| Profit Difference ($) | 376,063 | 1,625,548 |
| Profit Difference (% of Revenue) | 2.63% | 10.15% |

# Appendix

The probability of firm $i$ choosing action $k$ can be described as a function of state variables as follows (see equation 5 in the text):

$$P_i\left(a_i = k\right) = \int\int \mathbf{1}\left\{d_i\left(s, \epsilon_i^R, \epsilon_i^C\right) = k\right\} f\left(\epsilon_i^R, \epsilon_i^C\right) d\epsilon_i^R d\epsilon_i^C. \tag{13}$$

Recall as well that the revenue and cost equations are approximated as

$$R_i^k \approx R\left(s, a_i = k, a_{-i}; \theta_R^k\right) + \epsilon_i^R\left(k\right) + \eta_i^R\left(k\right),$$

$$C_i^k \approx C\left(s, a_i = k, a_{-i}; \theta_C^k\right) + \epsilon_i^C\left(k\right) + \eta_i^C\left(k\right). \tag{14}$$

Define,

$$\begin{aligned} \pi_i^k &= [R_i^k + \epsilon_i^R\left(k\right) + \eta_i^R\left(k\right)] - [C_i^k + \epsilon_i^C\left(k\right) + \eta_i^C\left(k\right)] && (15) \\ E\left(\pi_i^k\right) &= \overline{\pi}_i^k + \left[\epsilon_i^R\left(k\right) - \epsilon_i^C\left(k\right)\right] && (16) \\ \overline{\pi}_i^k &= R_i^k - C_i^k && (17) \end{aligned}$$

If strategy $k$ was chosen by firm $i$ (we ignore the market subscript in what follows) we know that

$$E\left[\pi_i^k\right] \geq E\left[\pi_i^{k'}\right] \quad \forall k \neq k'. \tag{18}$$

In other words

$$\overline{\pi}_i^k + \epsilon_i^R\left(k\right) - \epsilon_i^C\left(k\right) \geq \max_{k' \neq k}\left\{\overline{\pi}_i^{k'} + \epsilon_i^R\left(k'\right) - \epsilon_i^C\left(k'\right)\right\} \tag{19}$$

or

$$\begin{aligned} \epsilon_i^R\left(k\right) &\geq \max_{k' \neq k}\left\{\overline{\pi}_i^{k'} + \epsilon_i^R\left(k'\right) - \epsilon_i^C\left(k'\right)\right\} - \overline{\pi}_i^k + \epsilon_i^C\left(k\right) \\ \epsilon_i^R\left(k\right) &\geq \Delta\tilde{\pi}_i^k + \epsilon_i^C\left(k\right) && (20) \end{aligned}$$

where

$$\Delta\tilde{\pi}_i^k = \max_{k' \neq k}\left\{\overline{\pi}_i^{k'} + \epsilon_i^R\left(k'\right) - \epsilon_i^C\left(k'\right)\right\} - \overline{\pi}_i^k \tag{21}$$

Now,

$$E\left(\omega_i^R(k)\,|\epsilon_i^R(k)\geq\Delta\tilde{\pi}_i^k+\epsilon_i^C(k),\overline{\pi}_i\right)\neq 0 \tag{22}$$

In fact given the independence of $\epsilon$ and $\eta$,

$$E\left(\omega_i^R(k)\,|\epsilon_i^R(k)\geq\Delta\tilde{\pi}_i^k+\epsilon_i^C(k),\overline{\pi}_i\right)=E\left(\epsilon_i^R(k)\,|\epsilon_i^R(k)\geq\Delta\tilde{\pi}_i^k+\epsilon_i^C(k),\overline{\pi}_i\right). \tag{23}$$

Letting $g\left(\Delta\tilde{\pi}_i^k|\overline{\pi}_i\right)$ denote the density of $\Delta\tilde{\pi}_i^k$.this expectation can be written as

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\Delta\tilde{\pi}_i^k+\epsilon_i^C(k)}^{\infty}\frac{\epsilon_i^R(k)\,f\left(\epsilon_i^R(k),\Delta\tilde{\pi}_i^k,\epsilon_i^C(k)\,|\overline{\pi}_i\right)}{P\left(\epsilon_i^R(k)\geq\Delta\tilde{\pi}_i^k+\epsilon_i^C(k)\,|\overline{\pi}_i\right)}d\epsilon_i^R(k)\,d\Delta\tilde{\pi}_i^k d\epsilon_i^C(k)\\
=&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\Delta\tilde{\pi}_i^k+\epsilon_i^C(k)}^{\infty}\frac{\epsilon_i^R(k)\,f\left(\epsilon_i^R(k),\epsilon_i^C(k)\,|\overline{\pi}_i\right)g\left(\Delta\tilde{\pi}_i^k|\overline{\pi}_i\right)}{P\left(\epsilon_i^R(k)\geq\Delta\tilde{\pi}_i^k+\epsilon_i^C(k)\,|\overline{\pi}_i\right)}d\epsilon_i^R(k)\,d\Delta\tilde{\pi}_i^k d\epsilon_i^C(k)
\end{aligned}$$

Since by the i.i.d. assumption, $\epsilon_i^C(k)$ is independent of $\epsilon_i^R(k')$ and $\epsilon_{im}^C(k')$ .It is easy to see that this expectation will only be a function of profit indices, $\overline{\pi}_i=\left\{\overline{\pi}_i^1,\overline{\pi}_i^2,...,\overline{\pi}_i^K\right\}$.

In other words,

$$E\left(\omega_i^R(k)\,|\epsilon_i^R(k)\geq\Delta\tilde{\pi}_i^k+\Delta\epsilon_i^k+\epsilon_i^C(k),\overline{\pi}_i\right)=\Lambda_k(\overline{\pi}_i) \tag{24}$$

where $\Lambda_k$ is some function. Given the one to one correspondence between $\overline{\pi}_i$ and $\mathbf{P}_i$ this can equivalently be expressed as,

$$E\left(\omega_i^R(k)\,|\epsilon_i^R(k)\geq\Delta\tilde{\pi}_i^k+\Delta\epsilon_i^k+\epsilon_i^C(k),\overline{\pi}_i\right)=\Lambda_k(\mathbf{P}_i)$$

and the selectivity corrected regression can be run as,

$$R_i^k\left(\mathbf{s},\mathbf{a};\theta_R^k\right)=R\left(\mathbf{s},\mathbf{a};\theta_R^k\right)+\Lambda_k\left(\widehat{\mathbf{P}}_i\right)+\tilde{\omega}_i^R(k) \tag{25}$$

where $\Lambda_k(\mathbf{z})$ is some function of the vector $\mathbf{z}$, and $\widehat{\mathbf{P}}_i$ is a consistent estimator of $\mathbf{P}_i$ and $\tilde{\omega}_i^R(k)$ is a homoskedastic, mean zero error term. In practice the function $\Lambda_k$ can be approximated by available flexible methods (polynomial series, splines etc.)

# References

**Aguirregabiria, V. and Mira, P.**, 'Sequential Simulation Based Estimation of Dynamic Discrete Games', Forthcoming in *Econometrica* (2006).

**Ahn, H. and J.L. Powell** (1993) "Semiparametric Estimation of Censored Selection Models with a Nonparametric Selection Mechanism" *Journal of Econometrics* 58 pp. 3-29.

**Bajari, P., H. Hong, J. Krainer, and D. Nekipelov** (2005) "Estimating Static Models of Strategic Interactions", Working Paper, University of Michigan.

**Bell, D., Ho, T. , and Tang, C.S.,** "Determining Where to Shop: Fixed and Variable Costs of Shopping", *Journal of Marketing Research*, 35(3) (1998) pp. 352-369.

**Bell, D. and Lattin, J.** "Shopping Behavior and Consumer Preference for Store Price Format: Why 'Large Basket' Shoppers Prefer EDLP." *Marketing Science.* v.17-1 (1998) pp. 66-88

**Dahl, G.** (2002) "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets" *Econometrica* v. 70(6) pp. 2367-2420.

**Dubin, J. and D. McFadden** (1984) "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption" *Econometrica* 52 pp. 345-362.

**Ellickson, P and S. Misra** (2006) "Supermarket Pricing Strategies" Conditionally accepted at *Marketing Science*.

**Fox, J.** (2007). "Semiparametric Estimation of Multinomial Discrete Choice Models Using a Subset of Choices" Forthcoming in the *RAND Journal of Economics*

**Hanemann, W.M.** (1984) "Discrete/Continuous Models of Consumer Demand" *Econometrica* v. 52(3) pp. 541-561.

**Harsanyi, J.** (1973) "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points" *International Journal of Game Theory.* v. 2 pp. 1-23.

**Heckman, J.J.** (1979) "Sample Selection Bias as a Specification Error" *Econometrica* v. 47(1) pp. 153-161.

**Heckman, J.J. and B. Honore** (1990) "The Empirical Content of the Roy Model" *Econometrica* v. 58(5) pp. 1121-1149.

**Heckman, J.J. and R. Robb** (1985a) "Alternative Methods for Evaluating the Impact of Interventions" in *Longitudinal Analysis of Labor Market Data*, ed. J.J. Heckman and B.

Singer, pp. 156-245, Cambridge: Cambridge University Press.

**Heckman, J.J. and R. Robb** (1985a) "Alternative Identifying Assumptions in Econometric Models of Selection Bias" *Advances in Econometrics*, 5, pp. 243-87.

**Hoch, S., Dreze, X. and Purk, M.** "EDLP, Hi-Lo, and Margin Arithmetic" *Journal of Marketing*. v. 58 (1994) pp. 16-27.

**Hotz, J., and R. Miller** (1993) "Conditional Choice Probabilities and the Estimation of Dynamic Models", *Review of Economic Studies*, 60, pp. 497-531.

**Huff., D.L.** (1964) "Defining and Estimating a Trading Area" *Journal of Marketing* v. 28(3) pp. 34-38.

**Lal, R. and Rao, R.** "Supermarket Competition: The Case of Every Day Low Pricing." *Marketing Science*. v. 16-1 (1997) pp. 60-81.

**Lee, L.** (1982) "Some Approaches to the Correction of Selectivity Bias" *Review of Economic Studies* v. 49(3) pp. 355-372.

**Lee, L.** (1983) "Generalized Econometric Models with Selectivity" *Econometrica* v. 51(2) pp. 507-512.

**Mazzeo, M.** (2002) "Product Choice and Oligopoly Market Structure" *Rand Journal of Economics* v. 33(2), pp. 221-242.

**Mazzeo, M.** (2002) "Competitive Outcomes in Product-Differentiated Oligopoly" *Review of Economics and Statistics* v. 84(4), pp. 716-728.

**Pakes, A., J. Porter, K. Ho, and J. Ishii** (2006) "Moment Inequalities and Their Application " Working Paper: Harvard University.

**Reiss, P. and P. Spiller** (1989) "Competition and Entry in Small Airline Markets" *Journal of Law & Economics* v. 32(2), pp. S179-202

**Roy, A.D.** (1951) "Some Thoughts on the Distribution of Earnings" *Oxford Economic Papers* v. 3(2) pp. 135-146.

**Rust, J.** (1987) "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher" *Econometrica* 55-5 999-1033.

**Seim, K.** (2006) "An Empirical Model of Firm Entry with Endogenous Product-Type Choices" *Rand Journal of Economics*, 37(3) 619-640