Enriching interactions: Incorporating outcome data into static discrete games

Paul B. Ellickson · Sanjog Misra

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 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

P. B. Ellickson · S. Misra (✉)
William E. Simon School of Business Administration, University of Rochester, Rochester, NY 14627, USA
e-mail: misra@simon.rochester.edu

P. B. Ellickson
e-mail: paul.ellickson@simon.rochester.edu

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.

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 (2008) 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 will 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 classic techniques from the literature on selection (Heckman 1974, 1979; Heckman and Robb 1985, 1986; Ahn and Powell 1993) with recent two-step approaches to estimating discrete games (Aguirregabiria and Mira 2007; Bajari et al. 2010). The basic idea is quite simple. With data describing only discrete outcomes (e.g. entry, location, pricing strategy), we can recover the strategic parameters that govern profitability using existing methods adapted from the discrete choice literature. If, in addition, we also observe post-choice outcome data (e.g. revenues or costs), we can potentially decompose the determinants of profits into its more primitive components. However, to do so, we must first correct for selection bias stemming from the

fact that we observe only the revenues or costs for the action that was ultimately chosen. By definition, this is the choice that maximized profits relative to all other options, implying that any unobserved shock that influences this discrete choice must also have a direct impact on either revenue or costs (or both), inducing a selection bias that must be accounted for in the outcome equation. Fortunately, by framing the underlying game as one of incomplete information, flexible semi-parametric methods for selectivity correction can be applied directly. We can then use selectivity-corrected outcome equations to construct either revenue or cost counterfactuals (depending on which type of data are actually observed) that characterize the difference between the chosen action and any feasible alternative. Drawing on a revealed preference argument, a final inequality estimator can then be employed to recover the remaining parameters that govern either revenues or costs. In the simplest, fully parametric version of our approach, estimation consists of four steps, none more complicated that either a multinomial logit or a linear regression. Furthermore, because the initial stage of the estimation can be performed either market by market or firm by firm, our methods are robust to multiplicity of equilibria and can also be easily extended to account for heterogeneity, given access to sufficiently rich data.[1]

We illustrate our estimation technique with an empirical exercise that builds on an earlier analysis of supermarket pricing strategies. In a recent paper, Ellickson and Misra (2008) treat a supermarket's pricing decision as a discrete, store level game between rival chains and estimate the parameters of a reduced form profit function using techniques developed in Bajari et al. (2010). 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 choose their pricing strategy primarily to economize on costs. For example, firms that choose EDLP typically take a revenue hit relative to choosing a promotional strategy. However, the cost savings associated with EDLP outweigh the expected revenue loss, making it profitable on balance. Furthermore, we find that the *coordination* of strategies across firms is also driven by cost savings, rather than demand or revenue considerations. Having more rivals choose your same pricing strategy certainly reduces your revenues, consistent with a standard congestion or business-stealing story. However, it actually reduces your costs

---

[1] Multiplicity of equilibria can make it difficult to construct a likelihood since, in the absence of a clear selection rule, the model is effectively incomplete. If the researcher is willing to assume that only one equilibrium is played in the data, two-step methods restore completeness by allowing the (pseudo) likelihood function to condition on the equilibrium that was in fact selected. By estimating the first stage market by market, this equilibrium restriction can be weakened to requiring only that a single equilibrium be played in each market.

more. This suggests that, at least in the case of pricing strategies, the costs of differentiation can outweigh its benefits, leading to coordination.

Our approach draws on a rich body of literature spanning several areas of applied microeconomics and quantitative marketing. While empirical interest in discrete games of complete information dates to the seminal work of Bresnahan and Reiss (1991a, b), our approach is more closely connected to the incomplete information framework proposed by Rust (1994) and implemented by Seim (2006) and Bajari et al. (2010).[2] In particular, the incomplete information assumption is key to breaking a system of equations into a collection of single agent problems in which selection can be addressed directly. The connection between selectivity and discrete games was first addressed by Reiss and Spiller (1989), and later by Mazzeo (2002), both in the context complete information games. Zhu et al. (2009) use Mazzeo's complete information approach to incorporate revenue data into an entry model applied to the discount store industry. Our approach to selectivity clearly draws on the classic work of Heckman (1974, 1979), and most directly on the semi-parametric propensity-score based methods introduced by Heckman and Robb (1985, 1986) and Ahn and Powell (1993). The mixed continuous-discrete choice structure of the problem has obvious antecedents in both the consumer demand literature (Dubin and McFadden 1984; Hanemann 1984) and empirical applications of the (Roy 1951) model (Heckman and Honore 1990; Dahl 2002). Finally, the emphasis on inequality conditions derived from revealed preference arguments and the specific decomposition of the errors into structural and non-structural components draws upon the recent moment inequality literature (Pakes et al. 2006).

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 application and Section 4 concludes.

2 A model of strategic interactions

Our model of strategic interactions is based on the static discrete game framework analyzed in Bajari et al. (2010). After introducing the basic notation, profit maximization conditions, and equilibrium concept, we outline our four-step estimation procedure. Our algorithm proceeds as follows. The first step, which is common to the literature on discrete games, involves flexibly estimating firm beliefs using standard non- or semi-parametric techniques. Note that these conditional choice probabilities (CCPs) will also form the basis of the selection correction utilized in step two. The second step involves estimating

---

[2]There is a wide and growing literature on discrete games in both economics and marketing. Notable examples include Aguirregabiria and Mira (2007), Bajari et al. (2007), Berry (1992), Pakes et al. (2007), Draganska et al. (2009), Hartmann (2010), Ho (2009), Orhun (2006), Pesendorfer and Schmidt-Dengler (2007), Sweeting (2009), Vitorino (2007), and Zhu and Singh (2009). Ellickson and Misra (2011) provide an overview of this rapidly expanding field.

selectivity-corrected outcome equations, which yield consistent estimates of revenues (or costs, depending on the nature of available data) as a function of exogenous covariates and the (endogenous) actions of each firm's rivals. These revenue (or cost) functions represent one component of the payoffs each firm expects to receive from its discrete choice of action. To recover the second component, we turn to a revealed preference argument. Using the revenue (cost) functions, we can construct the revenues (costs) that the firm expected to earn (incur) from both the choice it actually made, as well as the remaining alternatives that it did not. If the chosen action did not yield the highest revenue (lowest cost), then it must have been driven by offsetting cost (revenue) advantages. Thus, the third step requires constructing counterfactual revenue (or cost) differences, which capture the first component of this trade-off. The fourth and final step uses inequalities based on the necessary conditions for profit maximization to recover the second component. Standard errors are obtained via the bootstrap.

2.1 A discrete game of incomplete information

Although we will introduce a more complex notation for the subsequent empirical exercise, we work here with a simplified structure that closely follows that of Bajari et al. (2010). 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 given by $\mathbf{s} = (s_1, ..., s_n) \in \Pi_i S_i$. 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 are assumed to be determined exogenously. For each firm, there are also two privately observed state variables. While each firm observes its own private state variables, they are known to the econometrician and rival firms only in distribution. These privately observed state variables are denoted $\epsilon_i^R(a_i)$ and $\epsilon_i^C(a_i)$, or more compactly ϵ_i^R and ϵ_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 ϵ_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 ϵ's are assumed to be *i.i.d.* across firms and actions. They are drawn from a joint distribution $f\left(\epsilon_i^R, \epsilon_i^C\right)$ will full support on $\mathbf{R}^2$, that is known to everyone, including the econometrician. Revenues and costs are also subject to a second pair of firm specific shocks denoted η_i^R and η_i^C (or, more compactly, η_i), each of which is assumed to be mean zero. Note that, in contrast to the private information components (ϵ_i^R and ϵ_i^C), this additional source of randomness is not observed by the firm prior to making its discrete choice over actions.

Following Pakes et al. (2006), these ex post realizations could represent either expectation errors (due, perhaps, to incomplete or asymmetric information) or errors in the measurement of revenues or costs. The key distinction is that, unlike ϵ_i^R and ϵ_i^C, η_i^R and η_i^C do not impact firm behavior, so they can be ignored when constructing beliefs. For notational simplicity, we will emphasize the "expectational" interpretation in what follows, but note that this distinction has no operational impact.

Firms choose strategies with the objective of maximizing expected profits

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

broken out here as revenue minus costs. More generally, we could work instead with *variable profits* minus *fixed* costs, which would allow access to a deeper level of structure. Since reliable data on costs (or margins) is rarely available, we emphasize the current structure here.

The revenue and cost terms are assumed to be structured as follows

$$R_i^k = R \left(\mathbf{s}, a_i = k, a_{-i}; \theta_R^k \right) + \epsilon_i^R \left(k \right) + \eta_i^R \left(k \right) \tag{2}$$

$$C_i^k = C \left(\mathbf{s}, a_i = k, a_{-i}; \theta_C^k \right) + \epsilon_i^C \left(k \right) + \eta_i^C \left(k \right), \tag{3}$$

where the deterministic components (R and C) need only belong to a finite-dimensional parametric family. Following the standard selectivity literature, we will assume additive separability of the stochastic components throughout. Additive separability is a standard assumption in the discrete games literature as well. Note that, while this structure is quite flexible, it does place some restrictions on the types of outcome data that can be accommodated by our approach.[3]

Consistent with the informational structure of the game, we assume firms choose the action a_i that yields the highest expected profit relative to all alternatives, so that

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

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

2.2 Constructing firm beliefs

As is now well known in the empirical games literature, the structure of a discrete game differs from a standard discrete choice model in that the actions of a firm's rivals enter its payoff function directly. In particular, the game structure transforms the standard single agent discrete choice problem into

---

[3]For example, if the researcher had access to detailed price and quantity data and chose to specify a discrete choice demand system with a "Nash in prices" supply side (e.g. Berry et al. 1995), the relevant structural errors would most naturally enter in a highly non-linear manner that could not easily be accommodated here.

a system of simultaneous discrete choice problems. Since this is a game of incomplete information, firms cannot perfectly predict the actions of their rivals. Moreover, since the ϵ's are treated as private information, a particular firm's decision rule $a_i = d_i\left(\mathbf{s}, \epsilon_i^R, \epsilon_i^C, E(\eta_i)\right) = d_i\left(\mathbf{s}, \epsilon_i^R, \epsilon_i^C\right)$ is a function of the state vector (**s**) and its *own* private information, but *not* the private information of its rivals.[4] In other words, each firm's beliefs over what its competitors will do will be a function 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

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

where $\mathbf{1}\left\{d_i\left(\mathbf{s}, \epsilon_i^R, \epsilon_i^C\right) = k\right\}$ is an indicator function equal to 1 if firm i chooses action k and 0 otherwise. We let **P** denote the vector of these probabilities. Since the firm does not observe the actions of its competitors prior to choosing its own action, its decision is based on these expectations.

2.3 Incorporating information on post choice outcomes

The inclusion of post-choice outcome data distinguishes our approach from existing two-step methods for estimating discrete games. While this auxiliary data could potentially take many forms (e.g. prices, quantities, revenues, costs, etc.), we focus on revenues since i) revenue data is often easier to obtain than information on costs, and ii) price and quantity data would require the researcher to impose additional structure (e.g. a demand system and a price setting mechanism) in which the structural errors would be unlikely to enter in an additively separable manner.[5] The inclusion of revenue data also matches the empirical example presented below. Clearly, the case where costs are observed instead is completely analogous (albeit less cleanly identified, as discussed below). Of course, having actual variable profit (or margin) data is also easily accommodated in our framework.

2.3.1 Calibrating revenue

Given data on revenues, recall that we have assumed that they can be decomposed as some function of the state vector and observed player actions, along with an additive (private information) revenue shock $\left(\epsilon_i^R(k)\right)$ and an additive expectation error $\left(\eta_i^R(k)\right)$

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

---

[4]Note that, because we have assumed that the expectational errors do not impact firm behavior, they drop out of the relevant decision rules.

[5]The extension of selectivity-correction techniques to settings with non-separable errors is beyond the scope of this paper.

Since the error components (η and ϵ) are indistinguishable from the perspective of the econometrician, the revenue function can equivalently be written as

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

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: firms choose the strategy that maximizes profits, based in part on the same unobservables $\epsilon_i^R(k)$ that impact revenues. In particular, $\mathcal{E}\left(\omega_i^R | a_i = k\right) \neq 0$, although the selectivity bias is clearly driven by $\epsilon_i^R(k)$. Consequently, consistent estimation of the revenue equation requires accounting for the sample selection bias induced by the underlying discrete game. Fortunately, the incomplete information structure allows this correction to take place using a standard, propensity-score based control function approach. We will specify how this is done below. However, before doing so, we must first parameterize the cost function.

2.3.2 Cost parameterization

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

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

Once again, ϵ is treated as private information while η is expectation error that is realized ex post. However, since costs are unobserved, we must now employ a latent variable approach and work with cost *differences* rather than *levels*. We adopt the following structure for these differences

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

where $\Delta\epsilon_i^C\left(k, k'\right) = \epsilon_i^C(k) - \epsilon_i^C(k')$ and

$$\Delta C_i\left(k, k' | s, \mathbf{P}, \theta_C\right) = \mathcal{E}\left(C\left(\mathbf{s}, a_i = k, a_{-i}; \theta_C^k\right) - C\left(\mathbf{s}, a_i = k', a_{-i}; \theta_C^{k'}\right)\right) \tag{10}$$

Note that since the firm makes its discrete decisions based on *expected* revenues and costs, the expectation errors (η) drop out, allowing realized actions to be replaced by expected actions (probabilities). In particular, because these "expectational" errors are unknown to the firm at the time that decisions are made, they 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 $\boldsymbol{\Psi}_i^k$ be an indicator function indicating that firm i chooses action k, conditional on the common parameter vector (θ), the common vector of state

variables (**s**), the firm's privately observed cost and revenue shocks (ϵ), and the expectations it has regarding the actions of its rivals (**P**). This can be written as

$$\Psi_i^k\left(a_i = k | \mathbf{P}, \mathbf{s}, \epsilon_i^R, \epsilon_i^C; \theta\right)$$

$$= \sum_{k' \neq k} \mathbf{1}\left(\mathcal{E}_\eta \left[\begin{array}{c} \pi_i\left(\mathbf{s}, a_i = k, \mathbf{P}_{-i}, \epsilon_i(k), \eta_i(k); \theta^k\right) \\ -\pi_i\left(\mathbf{s}, a_i = k', \mathbf{P}_{-i}, \epsilon_i(k'), \eta_i(k'); \theta^{k'}\right)\end{array}\right] > 0\right). \quad (11)$$

Let $g\left(\mathbf{r} | \mathbf{s}, \mathbf{a} = \mathbf{k}^*, \epsilon^R; \theta_R\right)$ be the joint conditional density of observed revenues (**r**) generated by the expectation error terms (η). The conditioning reflects the fact that the observed vector of revenues (across players), **r**, is conditioned on the state variables (**s**), the actions[6] of all players (**a**), the (private) revenue shocks $\left(\epsilon^R\right)$, and the revenue function parameters (θ_R). Finally, let $\mathbf{k}^*$ denote the vector of observed actions.

Given these components, the likelihood contribution from an arbitrary market can be written as

$$\mathbb{L} = \int g\left(\mathbf{r} | \mathbf{s}, \mathbf{a} = \mathbf{k}^*, \epsilon^R; \theta_R\right) f\left(\epsilon^R | \mathbf{a} = \mathbf{k}^*\right) d\epsilon^R$$

$$\times \prod_i \int \int \Psi_i^{k^*}\left(a_i = k^* | \mathbf{P}, \mathbf{s}, \epsilon_i^R, \epsilon_i^C; \theta\right) f\left(\epsilon_i^R, \epsilon_i^C\right) d\epsilon_i^R d\epsilon_i^C$$

where $f\left(\epsilon^R | \mathbf{a} = \mathbf{k}^*\right)$ is the marginal density of ϵ^R conditional on the observed vector of actions, while **P** has elements defined by

$$P_i^k\left(a_i = k\right) = \int \int \Psi_i^k\left(a_i = k | \mathbf{P}, \mathbf{s}, \epsilon_i^R, \epsilon_i^C; \theta\right) f\left(\epsilon_i^R, \epsilon_i^C\right) d\epsilon_i^R d\epsilon_i^C.$$

A few comments on the structure of the likelihood function are warranted here. First, note that ϵ^R and θ_R enter both g and Ψ, reflecting the underlying structural connection between the discrete choice of action and the resulting equilibrium outcome. This is similar to the discrete-continuous demand model proposed by Hanemann (1984), in which the choice of product impacts the choice of how much to consume through the underlying utility-theoretic framework. However, in contrast to Hanemann's consumer demand context, our setting involves multiple agents interacting in a discrete game, with the observed choice probabilities representing the equilibrium outcome of the game. The equilibrium constraint forces the agents expectations to match the corresponding equilibrium probabilities. This equilibrium constraint can complicate the evaluation of the likelihood function since, absent a clear selection rule, the direct nested fixed point approach (Rust 1987) to satisfying this constraint will be mis-specified if there is more than one fixed point (see Su and Judd 2007). Finally, we note that even with an equilibrium selection rule in place, maximizing the constrained likelihood is a numerically challenging

---

[6]Note that the *observed* revenues of a given store are a function of the *realized* actions of competing stores, not their *expected* actions. The expected actions of a store's rivals impact its *expected* revenues and, through those, its own action.

task. In what follows, we present a sequential estimation algorithm that greatly simplifies estimation.

2.4 Methodology: sequential estimation

In this section, we present our general estimation strategy, which involves four sequential steps. The intuition behind the algorithm is straightforward. Profits are the linear difference between revenues and costs. If we can characterize the impact of the state variables on revenues and use these functions to construct counterfactual revenues for the choices we do not observe, the conditions for profit maximization will capture the effect of the state variables on costs via revealed preference. The four steps that constitute our estimation algorithm are listed below. In what follows, we discuss each step in detail.

Step 1: Estimate each firm's beliefs via conditional choice probabilities (CCPs)
Step 2: Using these CCPs, estimate selectivity-corrected revenue equations
Step 3: Using these revenue equations, construct counterfactual revenue differences
Step 4: Estimate the remaining cost parameters from the profit inequalities

Step 1: Recovering equilibrium beliefs and constructing control functions The first step is exactly the same exercise employed in existing two-step approaches to estimating static games (e.g. Bajari et al. 2010). Note that here (as there) it is crucial that the first step provide consistent estimates $\widehat{\mathbf{P}}$ of the conditional choice probabilities, since these will in turn be used to construct estimates of players' expectations over rivals' actions. Furthermore, $\widehat{\mathbf{P}}$ is also a key component in the selectivity correction that follows, since we employ the control function approach suggested by Heckman and Robb (1985, 1986) and Ahn and Powell (1993). Clearly, there are numerous non- and semi-parametric approaches to constructing $\widehat{\mathbf{P}}$ in a flexible manner. These might include simple frequencies, the method of sieves, kernel or local linear regressions, or series estimators. Ideally, the researcher should use non-parametric methods to implement this step since the true functional form of these reduced form choice probabilities can rarely be specified a priori (even if the underlying profit function has a simple structure). Of course, if data limitations exist, one may have to employ less data intensive semi-parametric or even fully parametric methods. In such cases, it might be best to use an iterative approach (e.g. Aguirregabiria and Mira's (2007) Nested Pseudo Likelihood algorithm) to obtain consistent estimates of the choice probabilities. Note that an iterative approach could also be used to accommodate unobserved state variables, as in Ellickson and Misra (2008).

Step 2: Estimating selectivity-corrected revenue functions The second step recovers the parameters of the revenue function, correcting for the sample selection bias induced by the underlying discrete game. The approach advocated

here follows the semi-parametric, propensity-score based methods suggested by Heckman and Robb (1985, 1986) and Ahn and Powell (1993), which impose relatively weak assumptions on the underlying selection mechanism. Implementation is straightforward. We show in the "Appendix" that our approach amounts to running separate revenue regressions for each strategy (k), including the control function $\Lambda_k\left(\widehat{\mathbf{P}}_i\right)$, which is a flexible function of $\widehat{\mathbf{P}}_i$. Specifically, the regression equations are given by

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

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 this 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 (apart from the index restriction common to all of these methods). While this flexibility does come at some cost, the alternative would be to impose particular distributions on the ϵ and η 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 classic selectivity approach in which the errors in the two equations are simply assumed to be correlated (see e.g. Mazzeo 2002).[7]

We should also note here that our problem simplifies greatly due to the private information assumptions employed 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 actions enter the regression equation does not raise endogeneity issues. In particular, the private information assumption (on ε) allows the joint selectivity problem (that the revenues of *all* players are conditioned on the actions of *all* players) to be decomposed into a collection of individual selectivity problems.[8] The revenues may ex-post be correlated across players on account of the expectation errors (η's), but this raises no further difficulties.

---

[7]Note that this does not preclude the econometrician from employing specific distributions (e.g. extreme value errors for ϵ) since these assumptions can be imposed in the final step of our algorithm. Remaining agnostic about these errors at this point simply retains flexibility while also being fairly simple to implement.

[8]Note that in cases where the cardinality of an individual firm's choice set is large (i.e. there are many potential discrete actions), the researcher will likely face a dimensionality issue in modeling $\Lambda_k\left(\widehat{\mathbf{P}}_i\right)$ that is analogous to the curse of dimensionality associated with many multinomial choice problems. For example, constructing a control function via a second order polynomial approximation with J alternatives would require estimating $J + \Sigma_{i=1}^{J} i$ terms (so with 5 alternatives, one would have to estimate 20 parameters for the selectivity correction component alone). One possible solution is to assume that Dahl's (2002) *index sufficiency* assumption holds, and rely only on $\Lambda_k\left(\widehat{P_{iq}^k} : q \subset K\right)$. Unfortunately, this reduction in dimension is somewhat ad-hoc in that it is not based on any utility theoretic primitive and is inconsistent with many canonical examples (e.g. multinomial probit).

Step 3: Constructing counterfactual revenue differences Having obtained $\widehat{\theta}_R$, a consistent estimate of the parameter vector θ_R characterizing the revenue function, in step 2, the next step is to construct counterfactual revenue differences $\widehat{\Delta r_i^{kk'}} = \widehat{r}_i^k - \widehat{r}_i^{k'}$, computed from the difference between the fitted revenues constructed in step 1. These counterfactual revenue differences (CRDs) capture the difference in revenues associated with making choice k instead of k'. In practice, the CRDs are simply the fitted revenue differences between the strategy that was chosen (k) and those that were not (k'). Note here that $\widehat{\Delta r_i^{kk'}}$ 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 CRDs directly into the structural choice problem to estimate the cost differences without worrying about the private information components.[9]

Step 4: Estimating profit inequalities Once the CRDs are known, the conditions (4) ensuring that expected profits are maximized can be used to estimate the parameters of the implied cost differences. These necessary conditions yield the following empirical profit inequalities:

$$\widehat{\Delta r_i^{kk'}} - \Delta C_i\left(k, k'|\mathbf{s}, \widehat{\mathbf{P}}, \theta_C\right) \geq \kappa_i \tag{13}$$

where $\kappa_i = \Delta \epsilon_i^C - \Delta \epsilon_i^R$ represents the difference in the private information components for strategies k and k' and $\Delta C_i\left(k, k'|\mathbf{s}, \widehat{\mathbf{P}}, \theta_C\right)$ the cost difference defined by Eq. 10. Any set of parameters $\widehat{\theta}_C$ that satisfy Eq. 13 yield a consistent estimate of θ_C. Estimation can be performed parametrically (using a multinomial logit, for example) by imposing specific assumptions on κ (via appropriate assumptions on ϵ) or semi-parametrically using a maximum score approach. Alternatively, a bounds estimator similar to the moment inequalities methods proposed in Pakes et al. (2006) could be used instead. The estimation approach outlined here has been kept deliberately general. The actual implementation of the approach would require application-specific decisions on the part of the researcher. In the empirical application that follows, we provide an concrete example of how these choices might be made in practice.

2.5 Discussion

2.5.1 Identification

In empirical static discrete games, as in standard discrete choice models, identification of the latent profit or payoff parameters relies on the covariation

---

[9]In other words, simply plugging in revenue *data* for the observed choice would lead to mis-specification since these data include realizations of the private information components and measurement error shocks. The latter is explicitly *not* in the firm's information set when the discrete choice is being made.

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 should ideally 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 their expected actions). Given these exclusions, the strategic effects are identified since the variation in the beliefs a firm has over its competitors' actions are now driven at least in part by covariates that do not shift its own profits directly.[10] 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 in the second step regression equation poses identification problems. The solution in the selectivity literature (see e.g. Vella 1998) is to have a set of variables that impact the discrete choice but are excluded from the regression specification. 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 shift 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. In the context of retail competition, reasonable candidates might include land rent or property values, various zoning restrictions, or individual distances to distribution centers or central headquarters. Note that the two sets of exclusion restrictions (within and across firms) could overlap as long as they include some elements that are not common to both.

2.5.2 Inference

Since the multi-step structure of our estimation routine renders standard analytic approaches to inference intractable, 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 straightforward 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 follow directly from the arguments laid out in the selectivity (Ahn and Powell 1993) and static games (Bajari et al. 2010) literature. Because our algorithm is simply an extension of standard "two-step" procedures, the usual arguments based on Newey and McFadden

---

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

(1994) apply here as well. In general, as long as the first stage is consistently estimated, each subsequent stage will yield consistent estimates.

2.5.3 Sampling properties

To be clear, while the recipe outlined in this paper provides consistent estimates it is nevertheless inefficient relative to a full-solution approach as proposed by Rust (1987). In particular, the degree to which the first stage is estimated with noise directly impacts the amount of small sample bias generated in the estimator. Our estimator is not alone in this regard, as a similar criticism applies to two-step estimation of single agent dynamic problems as well as dynamic games, and two-step methods in general. Monte Carlo evidence examining this issue is available in a number of papers, including Hotz et al. (1994) and Aguirregabiria and Mira (2002, 2007). However, as we noted earlier our proposed estimator may be iterated to mitigate this bias by iterating on the fixed point mapping in probability space.[11] Aguirregabiria and Mira (2002) provide Monte Carlo evidence documenting both the small sample bias and the resulting improvements associated with this iteration procedure.

The actual magnitude of the small-sample bias depends on a number of factors including sample size, the dimensionality of the state space, the number of players, the cardinality of the action space and the flexibility afforded to the model in the first stage. While we cannot make a general statement as to the small sample properties of our estimator we conducted a limited Monte-Carlo investigation of its properties to assess its coverage properties.[12] In particular, we examined the role of sample size, number of players and the number of covariates with the rest of the specification fixed to those in our empirical application. Our overall results were quite encouraging: First, in no simulation did the coverage probability drop below 83% with the vast majority of simulations exhibiting decent coverage (90% or higher). In all cases, low coverage (<90%) occurred when the sample size *and* the number of players happened to be small relative to the number of covariates. Again, these coverage properties are specific to the particular assumptions of our game structure and we are not claiming them to be universally representative. To a researcher looking to use these methods, we would suggest undertaking a similar exercise in their particular application setting.

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

---

[11]This strategy implicitly requires that the equilibria that is played be stable. Pesendorfer and Schmidt-Dengler (2010) provide an example in which the equilibria is not stable and such iteration yields inconsistent estimates.

[12]Further details are available from the authors upon request.

([2008](#)) to incorporate store-level data on revenue. For the sake of brevity, we 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. ([2010](#)) and Aguirregabiria and Mira ([2007](#)), Ellickson and Misra ([2008](#)) find that competing stores coordinate on the choice of pricing 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 uncover 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 firm's beliefs, as well as 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 full set of bivariate interactions. Note that this allows for the possibility that a different equilibrium is played in each market.

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 control functions is not a concern. We do not need to make any additional assumptions (e.g. index sufficiency) to reduce the dimensionality. Given the large amount 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 and B-splines but found only very small differences in the resulting estimates.

The third step is straightforward, simply requiring the construction of predicted actual and counterfactual revenues, based on the observed covariates and the results from 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 constructed in step 3. Here, we again take a semi-parametric approach, this time based on pair-wise comparisons between the selected choice and each of the unchosen alternatives. In particular, we use a smoothed pair-wise

maximum score procedure based on the methods developed in Fox (2007). The results of an alternative specification employing a simple multinomial logit framework (not reported) were broadly similar. All standard errors were obtained by bootstrapping over markets.

3.2 Results and discussion

3.2.1 Step 1: estimating beliefs

As noted above, the 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 Fig. 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 observed measures of rival actions (the share of rival stores choosing EDLP and PROMO, respectively). Note that, in the estimation of the revenue regressions, we use the actual choices of the competitors, as opposed to expectations. 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 the revenue regression estimates for the EDLP, HYBRID, and PROMO strategies, while the last two columns
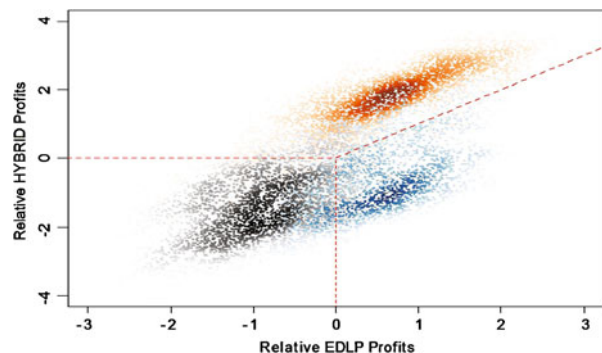
Fig. 1 Overall fit

Table 1 Selectivity corrected revenue regressions

	EDLP	HYBRID	PROMO	$\beta_E - \beta_P$	$\beta_H - \beta_P$
Intercept	−0.5072	−0.3148	0.3895	−0.8967	−0.7043
$\rho^E_{-i^{lm}_j}$	−0.0269	−0.0305	−0.0126	−0.0143	−0.0179
$\rho^P_{-i^{lm}_j}$	−0.0184	−0.0091^{ns}	−0.0220	0.0037	0.0129
Population density	0.3975	0.6347	0.3944	0.0031	0.2403
Store size	0.0803	0.0765	0.0696	0.0107	0.0069
Chain size	−0.0004	0.0001	0.0003	−0.0007	−0.0003
Proportion black	−0.4223	−0.4916	−0.3600	−0.0623	−0.1316
Proportion hispanic	0.5684	0.3848	−0.2475	0.8158	0.6323
Median HH income	0.0165	0.0214	0.0174	−0.0009	0.0040
Median HH # of vehicles	−0.5133	−0.5242	−0.4827	−0.0306	−0.0415
Median HH size	0.0027^{ns}	−0.1047^{ns}	0.1522	−0.1548	−0.2569

All coefficients are significant at the $\alpha = 0.05$ except where noted

contain the coefficients for EDLP and HYBRID differenced against those of PROMO. With just three exceptions, all coefficients are significant at the 5% level.

As one might expect, population density, store size, and median income all have a positive and significant impact on revenue, irrespective of the particular choice of pricing strategy. Similarly, the percentage of African American shoppers and the median number of vehicles have a negative impact. The latter covariate most likely reflects consumers' increased ability to search across stores, reducing the extent of local monopoly power and decreasing prices. 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" and thereby better able to mitigate the damage. There are three covariates that differ in their impact on revenues across the strategies. The percentage of Hispanic shoppers is associated with higher revenues for EDLP and HYBRID, but lower revenues for PROMO. Larger chains tend to generate greater revenues by choosing PROMO or HYBRID, rather than EDLP. Finally, the effect of household size is not significant under EDLP or HYBRID, but is positive and significant for PROMO.

Focusing on the strategic effects, we find that the revenues associated with every strategy are decreasing in competition, consistent with standard "business stealing" (congestion) arguments. Furthermore, the negative impact is largest for the share of rival firms who choose the same strategy as the focal firm, suggesting some scope for differentiation on the demand side. However, as we will see later, this revenue differential is small compared to the relative cost advantage of matching your rival's choice of pricing format.

Because these revenue regressions are key to backing out the implied cost differences, it is important to assess the robustness of our revenue specification to alternative functional forms. Since our focus is on the nature of strategic interaction, we introduce additional flexibility in the treatment of the remaining

controls and highlight the robustness of the strategic effects to these alterna-
tives. In particular, we implemented both a flexible semi-parametric model
based on splines and a factor analytic approach that incorporates a much
larger set of demographic variables (to control for the possible presence
of correlated unobservables). In the latter specification, 279 demographic
variables were factor analyzed to create thirty distinct factors that were then
used as composite market-level controls. Table 2 presents the results from
four specifications: a linear specification with no selection correction and
three alternatives that were selection corrected (our baseline linear model,
the semi-parametric spline specification, and the factor analytic approach).
A number of points are worth noting. First, the direction and magnitude
of the strategic effects are robust across the various specifications. Second,
ignoring the selectivity correction term seems to dampen the negative effect
of competitors employing the same strategy. Third, all strategic effects are
negative, suggesting that competitive effects are strong and verifying that the
differential effect of EDLP relative to PROMO ($\beta_E - \beta_P$) is negative across
specifications. Finally, the results from the specification with the composite
demographic factors exhibits similar strategic effects to the baseline model,
suggesting that our baseline specification is reasonably robust to correlated
unobservables.

3.2.3 Step 3: constructing counterfactual revenue differences

Having obtained revenue estimates in step 2, step 3 simply involves con-
structing counterfactual revenue differences: the expected gain (or loss) from
choosing an alternative pricing strategy. The results of this exercise are pre-
sented in Fig. 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 the stores that choose
EDLP would actually stand to *gain* revenue by switching to PROMO, implying
that substantial cost saving must be working to offset the forgone revenue

Table 2 Strategic effects in revenue regression: robustness

Model	EDLP		HYBRID		PROMO	
	$\rho^E_{-i^{lm}_j}$	$\rho^P_{-i^{lm}_j}$	$\rho^E_{-i^{lm}_j}$	$\rho^P_{-i^{lm}_j}$	$\rho^E_{-i^{lm}_j}$	$\rho^P_{-i^{lm}_j}$
No selectivity correction						
Linear	−0.0233	−0.0217	−0.0282	−0.0221	−0.0131	−0.0199
	(0.0056)	(0.0064)	(0.0074)	(0.0070)	(0.0066)	(0.0038)
Selectivity corrected models						
Linear	−0.0269	−0.0184	−0.0305	−0.0091	−0.0126	−0.0220
	(0.0056)	(0.0064)	(0.0074)	(0.0073)	(0.0067)	(0.0038)
Semiparametric	−0.0329	−0.0159	−0.0319	−0.0069	−0.0108	−0.0181
	(0.0055)	(0.0063)	(0.0073)	(0.0072)	(0.0064)	(0.0034)
Factor analytic	−0.0250	−0.0110	−0.0320	0.0022	−0.0080	−0.0177
	(0.0059)	(0.0064)	(0.0077)	(0.0073)	(0.0066)	(0.0037)

Fig. 2 Counterfactual
revenue differences in $MM
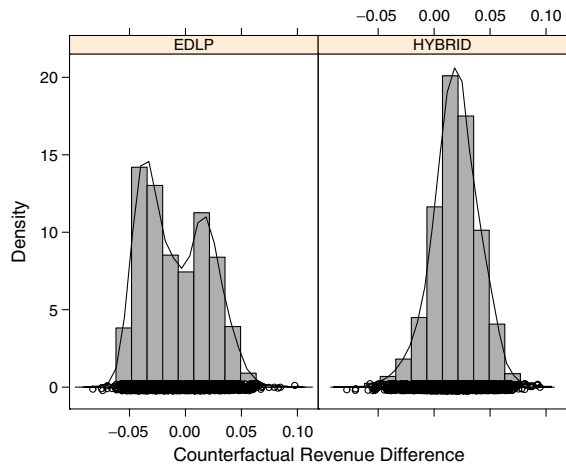
(a claim that we verify in the following section). Notably, the analogous data for HYBRID stores are quite different: the vast majority of HYBRID stores would actually *lose* revenue by switching to PROMO, indicating a more limited role for cost savings. However, to recover the actual cost implications 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 differences that rationalize their observed actions.

By examining the coefficient estimates, we can identify the covariates that drive these cost differences. Note that the estimates presented in Table 3 are directly comparable to those presented in the last two columns of Table 1, as everything is now expressed in differences. Since this is a reduced form, the demographic covariates are included as well, but don't have a clear interpretation. However, many firm characteristics enter as expected. For example, vertical integration and larger chain size reduce the costs of EDLP relative to PROMO, reflecting the firm level investments needed to profitably implement EDLP (greater scale is needed to smooth out the fluctuations in demand). Larger chain size also reduces the cost of HYBRID relative to PROMO, but vertical integration is insignificantly different from zero. Greater store size lowers the costs of PROMO relative to both alternatives, which may reflect the fact that a wider selection (as proxied by selling area) undercuts the benefits of smoothing demand (through a deeper, rather than wider, inventory). 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

Table 3 Cost estimates

Variable	EDLP		HYBRID	
	Estimate	Std. Err.	Estimate	Std. Err.
Intercept	1.1336	0.2962	−0.4636	0.2810
$\widehat{\rho}_{-i}^{E}$	−2.3243	0.2489	0.0491	0.2484
$\widehat{\rho}_{-i}^{P}$	2.3454	0.2599	5.7880	0.2573
Vertical integration	−0.2602	0.0782	0.0083	0.0690
Median rent	0.0001	0.0003	−0.0005	0.0003
Population density	0.0819	0.2771	−0.1870	0.2662
Store size	0.0217	0.0024	0.0030	0.0023
Chain size	−0.0013	0.0001	−0.0010	0.0001
Proportion black	−0.6212	0.1798	0.3406	0.1894
Proportion hispanic	0.4941	0.3602	−0.4670	0.3289
Median HH income	0.0038	0.0032	0.0043	0.0030
Median HH vehicles	−0.0931	0.1452	0.2031	0.1419
Median HH size	−0.2310	0.1060	−0.2644	0.1000

coordinating on EDLP actually reduces costs relative to PROMO. The same is true for HYBRID. Moreover, the magnitudes of these effects are significantly larger than what we found on the revenue side, indicating that the primary reason firms coordinate on pricing strategies is to reduce costs. We turn next to a broader discussion of these results.

3.3 Discussion

The empirical results from our estimation framework allow us to delve deeper into the mechanics of pricing format decisions. In particular, we are better able to grasp why and how stores choose amongst the various pricing strategies, and put a dollar value on the relative trade-offs. We discuss some of these new insights below.

Pricing Strategy Choice: With regard to pricing strategy, our findings reveal that the decision to implement EDLP is primarily driven by cost-side factors. We illustrate this point using a graphical analysis. Figure 3 plots both the revenue and cost differences of EDLP versus PROMO for stores that actually
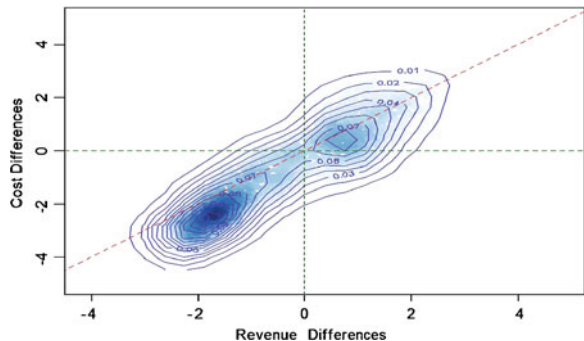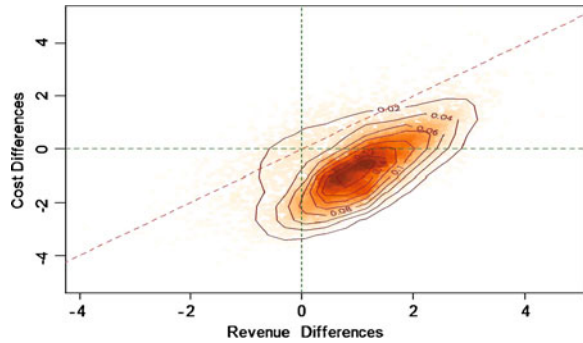
Fig. 3 EDLP versus PROMO

chose EDLP. The hashed red line represents the locus of indifference points
where the revenue and cost differences coincide. Note that the region with
the highest density of observations is the lower left quadrant, to the right of
the red line. This implies that, relative to PROMO, most stores choose EDLP
to economize on costs. Specifically, for the typical store choosing EDLP over
PROMO, the amount they expected to gain in cost savings outweighed the
amount they expected to lose in revenues. In contrast, Fig. 4 reveals that the
decision to choose HYBRID (rather than PROMO) is driven by a combination
of *both* revenues and costs. The highest density region is now in the lower right
quadrant, implying dominance along both dimensions. Note that this does not
imply that HYBRID is *always* the preferred choice, but rather that in those
cases where HYBRID offers the highest overall payoff, it does so on account
of both revenue gains *and* costs savings.

Positive Strategic Interactions: By decomposing latent profits into revenue
and costs components, we are better able to understand how strategic inter-
actions influence the choice of pricing strategy. In earlier work, Ellickson and
Misra (2008) found that pricing strategies are strategic complements, implying
that firms tend to collocate in pricing strategy space. The results presented here
suggest that this tendency to collocate is not driven by demand-side factors.
In fact, the estimates of strategic interactions on the revenue side suggest
that revenues actually decrease when a firm is faced with a larger fraction of
competitors adopting their same strategy. In contrast, the cost estimates reveal
that total costs are decreasing in the share of rivals adopting the same pricing
strategy. This makes intuitive sense. Greater similarity to one's rivals is likely
to drive down revenues, through a decrease in product differentiation and
increased competition. This provides an incentive to differentiate. However,
differentiation is costly. Firms have to convince consumers that their strategy
is superior to those of their rivals, primarily via increased advertising. Fur-
thermore, co-location might also reduce transaction costs when dealing with
common suppliers or manufacturers (this is presumably more important for
the smaller independent firms that do no vertically integrate into distribution).

Table 4 Comparison of strategies

Metric (mean)	EDLP	HYBRID
Cost savings (%)	5.75%	4.20%
Cost savings ($)	823,392	672,100
Revenues ($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 r evenue)	2.63%	10.15%

On net, the costs of differentiation outweigh the benefits when it comes to pricing strategy.

Monetary Valuation of Strategies: An additional benefit of our estimation approach is that we are able to scale latent profits in monetary units. To see why we can do so, note that the counterfactual revenue functions are evaluated in dollar terms and enter the discrete decision problem as an offset term. Since the cost parameters are now evaluated relative to this revenue offset, they can be interpreted on the same (dollar) metric. This allows us to monetize the costs and benefits of strategy choices in a way that has hitherto been impossible.

To assess the relative dollar magnitude of the overall estimated effects, we report the results of some simple calculations in Table 4. The table presents 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 provides a novel approach for incorporating information from continuous outcomes into static discrete games of incomplete information. We present a simple four step estimation algorithm, along with an empirical application that illustrates a particular implementation and the type of novel insight that can be gained from our approach. By incorporating revenue data into the discrete game framework, we can more clearly understand the nature of pricing decisions. In particular, we are able to distinguish between demand and cost side explanations and frame the latent profit implications in monetary terms. The methodology proposed here can be readily extended to incorporate dynamics, as demonstrated by Ellickson et al. (2011).

This framework can easily be adapted to accommodate richer structure. Our empirical application provides an example where only revenue data is available, yielding an analysis that relies on reduced forms for revenues and costs. In some cases, the reduced form may be of primary interest (for

example, in determining the extent to which two distinct formats or channels compete). In other cases, the deeper structural parameters may be required (to perform counterfactuals, for example). With additional data on costs (or profit margins), our framework provides a direct approach for recovering these parameters, while controlling for selection. We hope that this paper facilitates future research in this direction.

Appendix

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

$$P_i\left(a_i = k | \mathbf{s}\right) = \int \int \mathbf{1}\left\{d_i\left(s, \epsilon_i^R, \epsilon_i^C\right) = k\right\} f\left(\epsilon_i^R, \epsilon_i^C\right) d\epsilon_i^R d\epsilon_i^C. \quad (14)$$

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

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

$$C_i^k \approx C\left(\mathbf{s}, a_i = k, a_{-i}; \theta_C^k\right) + \epsilon_i^C\left(k\right) + \eta_i^C\left(k\right). \quad (15)$$

Define,

$$\pi_i^k = \left[R_i^k + \epsilon_i^R\left(k\right) + \eta_i^R\left(k\right)\right] - \left[C_i^k + \epsilon_i^C\left(k\right) + \eta_i^C\left(k\right)\right] \quad (16)$$

$$E\left(\pi_i^k\right) = \bar{\pi}_i^k + \left[\epsilon_i^R\left(k\right) - \epsilon_i^C\left(k\right)\right] \quad (17)$$

$$\bar{\pi}_i^k = R_i^k - C_i^k \quad (18)$$

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

$$E\left[\pi_i^k\right] \geq E\left[\pi_i^{k'}\right] \quad \forall k \neq k'. \quad (19)$$

In other words

$$\bar{\pi}_i^k + \epsilon_i^R\left(k\right) - \epsilon_i^C\left(k\right) \geq \max_{k' \neq k}\left\{\bar{\pi}_i^{k'} + \epsilon_i^R\left(k'\right) - \epsilon_i^C\left(k'\right)\right\} \quad (20)$$

or

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

$$\epsilon_i^R\left(k\right) \geq \Delta\tilde{\pi}_i^k + \epsilon_i^C\left(k\right) \quad (21)$$

where

$$\Delta\tilde{\pi}_i^k = \max_{k' \neq k}\left\{\bar{\pi}_i^{k'} + \epsilon_i^R\left(k'\right) - \epsilon_i^C\left(k'\right)\right\} - \bar{\pi}_i^k \quad (22)$$

Recalling that $\omega_i^R\left(k\right) = \epsilon_i^R\left(k\right) + \eta_i^R\left(k\right)$, it is clear that

$$E\left(\omega_i^R\left(k\right) | \epsilon_i^R\left(k\right) \geq \Delta\tilde{\pi}_i^k + \epsilon_i^C\left(k\right), \bar{\pi}_i\right) \neq 0 \quad (23)$$

Of course, given the independence of ϵ and η,

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

Now, 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

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

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

$$\times\, d\epsilon_i^R(k)\, d\Delta\tilde{\pi}_i^k d\epsilon_i^C(k)$$

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,

$$E\left(\omega_i^R(k)\,|\epsilon_i^R(k) \geq \Delta\tilde{\pi}_i^k + \Delta\epsilon_i^k + \epsilon_i^C(k), \overline{\pi}_i\right) = \Lambda_k(\overline{\pi}_i) \tag{25}$$

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

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

where we have abused notation slightly in using $\Lambda(\cdot)$ to represent both functions. The selectivity corrected regression can then be run as

$$R_i^k(\mathbf{s}, \mathbf{a}; \theta_R^k) = R\left(\mathbf{s}, \mathbf{a}; \theta_R^k\right) + \Lambda_k\left(\widehat{\mathbf{P}}_i\right) + \tilde{\omega}_i^R(k) \tag{26}$$

where $\Lambda_k(\mathbf{z})$ is a 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 Λ_k can be approximated by standard methods (polynomial series, splines, etc.)

References

Aguirregabiria, V., & Mira, P. (2002). Swapping the nested fixed point algorithm: a class of estimators for discrete Markov decision models. *Econometrica, 70*(4), 1519–1543.

Aguirregabiria, V., & Mira, P. (2007). Sequential estimation of dynamic discrete games. *Econometrica, 75*(1), 1–53.

Ahn, H., & Powell, J. L. (1993). Semiparametric estimation of censored selection models with a nonparametric selection mechanism. *Journal of Econometrics, 58*, 3–29.

Bajari, P., Benkard, L., & Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica, 75*(5), 1331–1370.

Bajari, P., Hong, H., Krainer, J., & Nekipelov, D. (2010). Estimating static models of strategic interactions. *Journal of Economics and Business Statistics, 28*(4), 469–482.

Berry, S. (1992). Estimation of a model of entry in the airline industry. *Econometrica, 60*(4), 889–917.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica, 63*(4), 841–890.

Bresnahan, T., & Reiss, P. (1991a). Empirical models of discrete games. *Journal of Econometrics, 48*, 57–81.

Bresnahan, T., & Reiss, P. (1991b). Entry and competition in concentrated markets. *Journal of Political Economy, 99*, 977–1009.

Dahl, G. (2002). Mobility and the return to education: Testing a Roy model with multiple markets. *Econometrica, 70*(6), 2367–2420.

Draganska, M., Mazzeo, M., & Seim, K. (2009). Beyond plain vanilla: Modeling product assortment choices in the ice cream market. *Quantitative Marketing and Economics, 7*(2), 105–146.

Dubin, J., & McFadden, D. (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica, 52*, 345–362.

Ellickson, P., & Misra, S. (2008). Supermarket pricing strategies. *Marketing Science, 27*(5), 811–828.

Ellickson, P., & Misra, S. (2011). Estimating discrete games. *Marketing Science* (forthcoming).

Ellickson, P., Misra, S., & Nair, H. (2011). Repositioning dynamics and pricing strategy. Simon School of Business Working Paper No. FR 11-07.

Fox, J. (2007). Semiparametric estimation of multinomial discrete choice models using a subset of choices. *RAND Journal of Economics, 38*(4), 1002–1019.

Hanemann, W. M. (1984). Discrete/Continuous models of consumer demand. *Econometrica, 52*(3), 541–561.

Hartmann, W. R. (2010). Demand estimation with social interactions and the implications for targeted marketing. *Marketing Science, 29*(4), 585–601.

Heckman, J. J. (1974). Shadow prices, market wages, and labor supply. *Econometrica, 42*, 679–693.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47*(1), 153–161.

Heckman, J. J., & Honore, B. (1990). The empirical content of the Roy model. *Econometrica, 58*(5), 1121–1149.

Heckman, J. J., & Robb, R. (1985). Alternative methods for evaluating the impact of interventions. In J. J. Heckman, & B. Singer (Eds.), *Longitudinal analysis of labor market data* (pp. 156–245). Cambridge: Cambridge University Press.

Heckman, J. J., & Robb, R. (1986). Alternative methods for solving the problem of selection bias in evaluating the impact of treatments on outcomes. In H. Wainer (Ed.), *Drawing inferences from self-selected samples*. New York: Springer.

Ho, K. (2009). Insurer-provider networks in the medical care market. *American Economic Review, 99*(1), 393–430.

Hotz, V.J., Miller, R.A., Sanders, S., & Smith, J. (1994). A simulation estimator for dynamic models of discrete choice. *Review of Economic Studies, 61*(2), 265–289.

Lal, R., & Rao, R. (1997). Supermarket competition: The case of every day low pricing. *Marketing Science, 16*(1), 60–81.

Mazzeo, M. (2002). Competitive outcomes in product-differentiated oligopoly. *Review of Economics and Statistics, 84*(4), 716–728.

Newey, W. K., & McFadden, D. (1994). Large sample estimation and hypothesis testing. In D. McFadden, & R. Engle (Eds.), *Handbook of econometrics* (Vol. 4, pp. 2113–2245). North-Holland, Amsterdam, The Netherlands: Elsevier.

Orhun, Y. (2006). *Spatial differentiation in the supermarket industry*. Working paper, GSB Chicago.

Pakes, A., Ostrovsky, M., & Berry, S. (2007). Simple estimators for the parameters of discrete dynamic games (with entry-exit examples). *RAND Journal of Economics, 38*(2), 373–399.

Pakes, A., Porter, J., Ho, K., & Ishii, J. (2006). *Moment inequalities and their application*. Working paper, Harvard University.

Pesendorfer, M., & Schmidt-Dengler, P. (2007). Asymptotic least squares estimators for dynamic games. *Review of Economic Studies, 75*, 901–928.

Pesendorfer, M., & Schmidt-Dengler, P. (2010). Sequential estimation of dynamic discrete games: A comment. *Econometrica, 78*(2), 833–842.

Reiss, P., & Spiller, P. (1989). Competition and entry in small airline markets. *Journal of Law and Economics, 32*(2), S179–202.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers, 3*(2), 135–146.

Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica, 55*(5), 999–1033.

Rust, J. (1994). Estimation of dynamic structural models, problems and prospects: Discrete decision processes. In C. Sims (Ed.), *Advances in econometrics: Sixth world congress* (Vol. 2, Chap. 4, pp. 119–170). Cambridge: Cambridge University Press.

Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices. *RAND Journal of Economics, 37*(3), 619–640.

Su, C. L., & Judd, K. L. (2007). *Constrained optimization approaches to estimation of structural models*. Working paper, CMS-EMS, Kellogg School of Management.

Sweeting, A. (2009). The strategic timing of radio commercials: An empirical analysis using multiple equilibria. *RAND Journal of Economics, 40*(4), 710–742.

Vella, F. (1998). Estimating models with sample selection bias: A survey. *Journal of Human Resources, 33*, 127–172.

Vitorino, M. A. (2007). *Empirical entry games with complementarities: An application to the shopping center industry*. Working paper, Chicago GSB.

Zhu, T., & Singh, V. (2009). Spatial competition and endogenous location choices: An application to discount retailing. *Quantitative Marketing and Economics, 7*(1), 1–35.

Zhu, T., Singh, V., & Manuszak, M. (2009). Market structure and competition in the retail discount industry. *Journal of Marketing Research, 46*(4), 453–466.