Modelling Pricing Behavior with Weak A-Priori Information: Exploratory Approach

Carlo Russo and Massimo Sabbatini
University of Cassino, DIMET, Italy
russocar@unicas.it ; m.sabbatini@caspur.it

Abstract

In the absence of reliable a priori information, choosing the appropriate theoretical model to describe an industry's behavior is a critical issue for empirical studies about market power. A wrong choice may result in model misspecification and the conclusions of the empirical analysis may be driven by the wrong assumption about the behavioral model.

This paper develops a methodology aimed to reduce the risk of misspecification bias. The approach is based on the sequential application of a sliced inverse regression (SIR) and a nonparametric Nadaraya-Watson regression (NW). The SIR-NW algorithm identifies the factors affecting pricing behavior in an industry and provides a non-parametric characterization of the function linking these variables to price. This information may be used to guide the choice of the model specification for a parametric estimation of market power.

The SIR-NW algorithm is designed to complement the estimation of structural models of market behavior, rather than to replace it. The value of this methodology for empirical industrial organization studies lies in its data-driven approach that does not rely on prior knowledge of the industry. The method reverses the usual hypothesis-testing approach. Instead of first choosing the model based on a priori information and then testing if it is compatible with the data, the econometrician selects a theoretical model based on the observed data. Thus, the methodology is particularly suited for those cases where the researcher has no a priori information about the behavioral model, or little confidence in the information that is available .

1 An Overview of the Model-Specification Problem

Industrial organization has described imperfect competition using a broad set of theoretical models. Static vs. dynamic frameworks, price vs. quantity choice, and homogeneous vs. differentiated products are just a few of the modeling decisions that economists must make. The abundance of available theoretical models is a critical challenge for applied economists wishing to estimate market power using a structural approach. Each theoretical model implies a different econometric model, and the econometrician must be able to select the one corresponding to the "true" data-generating process. Model specification is an important issue because a wrong choice will cause biased estimates.

The empirical literature has addressed the model specification problem using three basic strategies: adopting general models nesting a broad set of potential behavioral models, using the NEIO approach, and testing multiple model specifications. In the first approach the researcher obtains information about the behavior of the industry as the result of the estimation of a general model: the values of the estimated parameters identify the behavioral model and measure the degree of competition simultaneously. The analysis of the rice export market by Karp and Perloff (1989) is an example of this approach. Their approach nests four models: collusion, price taking, Nash-Cournot open loop, and Nash-Cournot with feedback. Identification is based on the estimation of a "behavioral parameter" which can also account for "intermediate paths" (a similar approach was followed by Katchova, Sheldon and Miranda 2005). The methodology allows the authors to account explicitly for dynamic adjustments in industry behavior, avoiding a potential bias in the estimates (Karp and Perloff 1993; Slade 1995). The use of nesting models does not solve the model specification problem, since no framework may nest all possible alternatives. Thus, researchers are still forced to choose across alternative "families" of models, and a wrong decision still may lead to biased estimates. Moreover, the solution of the model may be quite complicated and may require the introduction of simplifying assumptions in order to obtain an explicit form of the structural equations for the econometric model. For example, Karp and Perloff (1989) impose a linear-quadratic form in order to estimate the feedback model. The imposition of these restrictions may lead to biased estimation of the model parameters, if they are not compatible with the true data-generating process.

Since the late 1980s, the NEIO approach has been one of the most popular frameworks for empirical analysis of market power (Sheldon and Sperling 2003). The approach assumes that the industry behavior can be summarized by a finite (and known) number of parameters. These parameters are able to describe completely the equilibrium of any unknown economic game (e. g., Appelbaum 1982; Dockner 1992). Thus, the econometrician can estimate the conduct parameters and obtain an estimate of the degree of market power without specifying the underlying behavioral model. The result of the empirical analysis can be interpreted as an "as if" measure of industry conduct. This agnostic approach allowed researchers to use the NEIO framework despite intensifying criticism of conjectural variation theory (e. g., Makowski 1987). Bresnahan (1989) emphasized the substantial difference between conjectural variations and NEIO as empirical tools. The former uses the conduct parameters to represents firms' expectations about what happens if they deviate from the collusive agreement, while the latter estimates the parameters to infer what firms do as a result of these expectation.

The NEIO approach does not solve the model specification problem because it does not ensure that the market power estimates are unbiased. As noted in the introduction, Corts (1999) argued that the application of the NEIO approach may lead to biased estimations of the firms' marginal costs and, consequently, of the market power parameters. The essence of Corts' critique is that the two-stage least square structural model used by NEIO captures only variations of the Lerner index at the margin ("equilibrium variations"), while the recovery of the conduct parameter requires the average variation ("equilibrium values"). Thus, the estimation is unbiased only when the marginal variation coincides with the average variation, as in the case of Cournot oligopoly. The general conclusion from Corts' critique is that "Without stipulating the true nature of the behavior underlying the observed equilibrium, no inference about the extent of market power can be made from analysis of the observed variables" (Corts 1999, p. 229).

The third strategy is based on testing the model specification explicitly. In general the approach is the following: the researcher estimates the model over a set of alternative functional forms and then selects the one that fits the data best, usually using a likelihood ratio (LR) test of some form. The literature in this field is remarkably sparse, considering the importance of the topic. Notable examples include the study of collusive behavior in the U.S. automobile industry by Bresnahan (1987) the analysis of competition in the carbonated soda industry by Gasmi, Laffont and Voung (1992) and the analysis of bilateral market power in the U.S. leaf tobacco market by Raper, Love and Shumway (2000).

Gasmi, Laffont and Vuong use a two-step approach. In the first step, they specify demand and cost functions for each firm in the market in order to characterize the payoff functions for the firms in the industry. Then they introduce a set of assumptions describing alternative forms of market behavior and calculate the corresponding "equilibrium paths". The first order conditions associated with these paths, together with the firms' demand functions, comprise the econometric model for each market conduct model. In other words, the authors estimate each market conduct model "as if" it were the true one, using a structural approach. In the second step, the authors used a likelihood-ratio test developed by Vuong (1989) in order to select the model that was most consistent with the observed data across all pair-wise comparisons of the estimated models' log-likelihood functions. For each pair-wise comparison, the null hypothesis is that the two competing models explain the data equally well and the alternative hypothesis is that one model fits better. Vuong's test is particularly suited to the problem at hand because it does not require that either model is correctly specified. Since Gasmi, Laffont and Vuong introduced the use of the procedure for testing market structure, Vuong's test has mainly been used to identify the nature of the competition in international markets (e. g., Carter and MacLaren 1997; Dong, Marsh and Stiegert 2006).

Raper, Love and Shumway used a similar approach to identify the most likely behavioral model in a broad set of models including perfect competition, monopoly, monopsony, duopoly, duopsony and cooperative bilateral monopoly. The main difference from the Gasmi, Laffont and Vuong paper is the nature of the criterion used for choosing the best model specifications. In this case, the authors used Pollak and Wales' likelihood dominance criterion (Pollak and Wales 1991). This approach is based on pair-wise comparisons across the model specifications. In each pair-wise comparison, one model is considered to dominate the other if the difference between the log-likelihood values is greater than a pre-specified threshold. Bresnahan's analysis of the 1955 automobile price war followed a similar structure, testing Bertand-Nash, collusive and hedonic price models in the U.S. automobile industry using a Cox test (Cox 1961).

The tests used by the three papers share two major limitations: they rely on specific assumptions regarding demand and cost functions, and they pick the best alternative among the given set, but this choice does not necessarily correspond to the true data-generating process. The reliance on specific functional forms implies that the general approach used in the three papers is actually a joint test on the behavioral model the distribution of the error term, and the functional form specifications (Nevo 2001). These methods require evaluating the likelihood of each model, which can be derived only after making non-trivial assumptions in addition to the specification of the behavioral model. Thus, a correct behavior specification may be rejected if these additional assumptions introduce a substantial bias.

The very nature of the model-specification approach creates a more fundamental problem for an econometrician who wishes to determine the nature of competition in a specific market. The tests are based on pair-wise comparisons of a set of *m* alternative model specifications. The preferred model is the one that fits the data best compared to the *m*-1 alternatives, but this does not necessarily imply that this selection corresponds to the true data-generating process. In other words, if the true economic behavior of the industry is not included in the *m* alternatives, the estimation of the market power parameter may still be biased. The choice of the set of model specifications is critical for obtaining unbiased estimates and the number of alternatives should be large, including a wide range of solution concepts and demand and cost functional forms. However, as the number of models increases, the pair-wise approach may become impractical.

The next section develops a new exploratory approach to the model specification problem. The methodology allows the researcher to identify the key variables of the industry's behavior, without imposing any a priori assumptions. This information can be used to select the model specification that is closest to the "true" data-generating mechanism.

2 Empirical methodology

Assume that we observe a set of variables that may or may not affect the pricing behavior of the industry and that we can divide the available information at time *t* into two matrices: a *T´S* matrix of exogenous variables (*X*) representing the shifters of demand, supply and marginal cost of processing, and a *T´*1 matrix of endogenous variables (*Y*) representing the price. Using a two-step approach, which we refer to as the SIR-NW algorithm, we identify the effects of these variables on the price.

The intuition behind our two-stage approach is simple. The obvious methodological approach to estimating how the exogenous variables affect the margin without imposing specific function forms is to use non-parametric regression techniques. Yet, if *S*, the number of exogenous regressors, is large, this approach is likely to suffer from the *curse of dimensionality*: adding extra dimensions to the regression space leads to an exponential increase in volume, which slows the rate of convergence of the estimator exponentially. In order to avoid this curse, we compress the original set of variables into a smaller number of factors that are linear combinations of the variables using Sliced Inverse Regression, a dimension reduction technique (Li, 1991).

More formally, the SIR-NW algorithm is implemented in two steps. The first step identifies the Sliced Inverse Regression factors, compressing the original set of variables into a smaller number of factors that are linear combinations of the variables. Chen and Smith (2007) showed that these factors can be used as non-parametric regressors, for example, in a Nadaraya –Watson (NW) estimation, which is the procedure we apply in the second step of our approach.

3 A Simulation

This section uses simulation in order to assess the ability of the exploratory approach to identify the industry's competitive behavior by applying the SIR-NW algorithm to three economic models: perfect competition, a symmetric Cournot oligopoly model with fixed number of firms (labeled as the Cournot model), and a Rotemberg-Saloner dynamic, collusive supergame (Rotemberg and Saloner 1986).

The data for the three models were generated using the following linear inverse demand and constant marginal cost functions:

$$
\begin{aligned}
P_t &= 100 + 20 \cdot D_t - Q_t + \varepsilon_t \\
MC_t &= 10 + W_t
\end{aligned}
\tag{1}
$$

where *P* is price, *Q* is the industry equilibrium quantity, *MC* is the marginal cost, which is assumed to be identical across firms, and *e* is an i.i.d. error term distributed as a standard normal. *D* and *W* are the demand and marginal cost shifters, respectively, and they are distributed according to a identical and independent discrete uniform distribution with $D=1,2,\dots,K_D$ and $W=1,2,\dots,K_W$.

Since *D* and *W* are independent, each state of nature has a probability of .

$$1/(K_D \cdot K_W)$$

In the simulation $K_D=K_W=10$.[1] The Cournot and Rotemberg-Saloner models maintain a fixed number of firms (*M*) with *M*=6.
The equilibrium quantity and price under perfect competition are calculated using the following equations:

$$\begin{gathered} Q_t^{PC} = 100 + 20 \cdot D_t - 10 - W_t \\ . \\ P_t^{PC} = 10 + W_t \end{gathered} \tag{2}$$

The equilibrium quantity and price in the Cournot model are:

$$\begin{gathered} Q_t^{C} = \frac{6 \cdot \left(100 + 20 \cdot D_t - 10 - W_t\right)}{7} \\ . \\ P_t^{C} = \frac{\left(100 + 20 \cdot D_t\right) + 6(10 + W_t)}{7} \end{gathered} \tag{3}$$

In the Rotemberg-Saloner model, the firms set their individual production levels so that their joint profit is maximized under the constraint that no firm has incentive to break the collusive agreement. Given equation (2) and *M*=6, the solution of the constrained maximization problem gives the following equilibrium quantity:

$$Q_t^{RS} = 6 \cdot \max\left(\tilde{q}_t, q_t^m\right)$$

with:

$$\begin{gathered} q_t^m = \frac{90 + 20D_t - W_t}{12}, \\ \tilde{q}_t = \frac{90 + 20D_t - W_t}{7} - \frac{2\sqrt{L}}{7} \end{gathered}$$

where *L* is the expected discounted loss of profit incurred by entry into the punishment phase in period *t*. Note that the punishment scheme is time-invariant and state-independent. In order to calculate the present value of the profit loss I used an interest rate of 50%, which corresponds to a discount factor of 0.667. While this discount factor may seem low, it was chosen to reflect Rotemberg and Saloner's (1986, p. 394) comment that infinite punishments seem "unrealistic" in the real world.
The equilibrium prices in the Rotemberg-Saloner model are:

1. The distribution of the shifter is consistent with Corts' (1999) modeling of the Rotemberg-Saloner model. Although Corts allowed for serially correlated demand shifts according to a Markov process, I restrict the analysis to the special case of i.i.d. demand shocks.

$$P_t^{RS} = 55 + 10D_t + 0.5W_t \qquad \text{if } Q_t^{RS} = 6 \cdot q_t^m$$
$$P_t^{RS} = \frac{(100 + 20D_t) + 6(10 + W_t)}{7} + \frac{12\sqrt{L}}{7} \qquad \text{otherwise.} \tag{4}$$

In each model an error term is added to the equilibrium quantity. It takes the form ,

$$e_t = 0.5 \cdot \varepsilon_t + 0.5 \cdot u_t$$

where *u* follows a standard normal distribution and *e* is the error term from equation (1). This structure of the error term ensures that quantity is endogenously determined with price.
The SIR-NW algorithm was applied to a data series generated by each of the three models. A SIR of the price data series on the *X* matrix identified the relevant factors (the *H* matrix), then a non-parametric regression of price on *H* estimated the linking function (F_0).

The matrix *X* was composed of four variables: *W*, *D*, *F* and *WN* (a white noise variable distributed as a standard normal). Each simulated data series included 500 observations, which were divided into 10 slices of fifty observations each. The critical regions of the statistical tests were set using a 99% confidence level. The non-parametric regressions were based on a Nadaraya-Watson kernel estimator and a Silverman bandwidth.

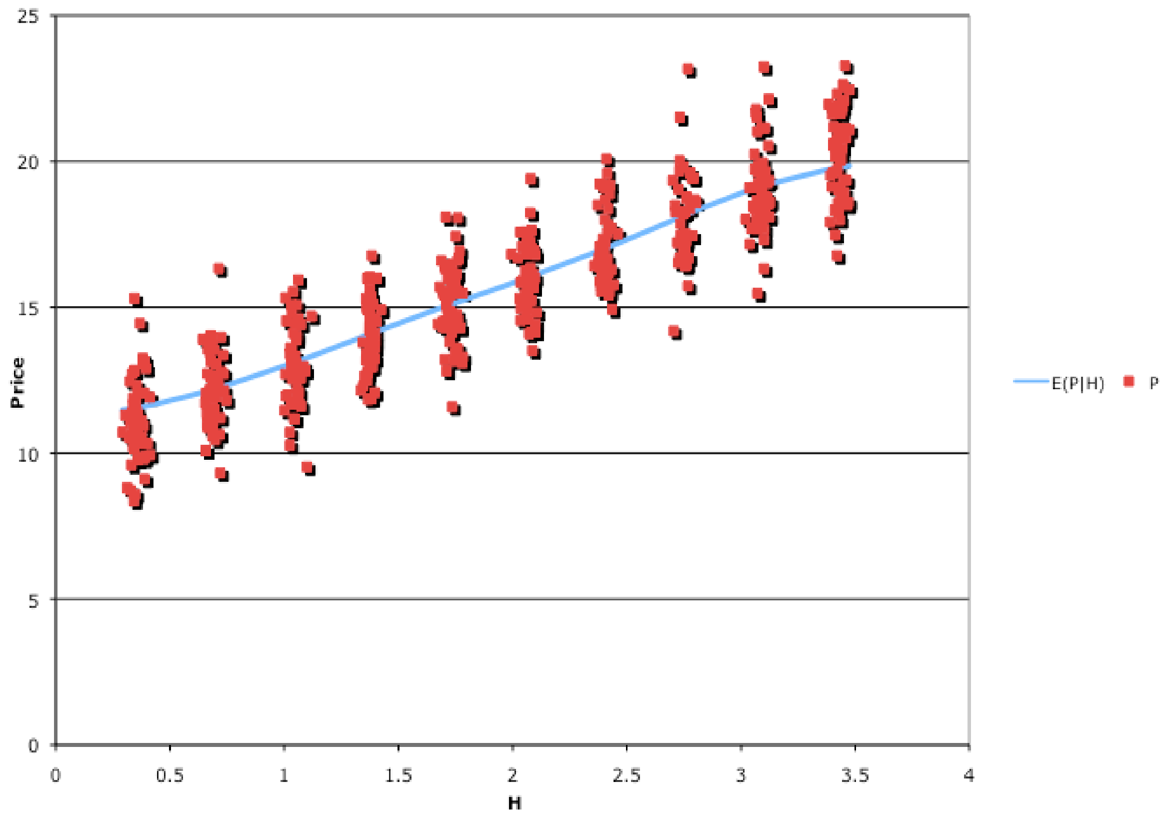

Figure 1. Non-parametric regression of price on the first SIR factor in a perfect competition model.

The SIR on the perfect competition data series identified one relevant factor defined by the following linear combination of the elements of the *X* matrix:

$$H_t = -0.003D_t - 0.341W_t - 0.023F_t + 0.003WN_t$$
$$(-0.410) \quad (-51.308) \quad (1.193) \quad (0.128)$$

where the numbers in parenthesis are the t-statistics on the beta coefficients. The SIR indicated that the only statistically significant variable affecting the pricing behavior of the industry was the marginal cost shifter. The non-parametric univariate regression of *P* on *H* suggested that the linking function is linear (Figure 2.1). These results were consistent with the pricing rule for the perfectly competitive model in equation .
The SIR analysis of the symmetric Cournot model identified two factors:

$$H_t^1 = -0.350D_t - 0.100W_t + 0.007F_t + 0.004WN_t$$
$$(-109.580) \quad (-30.610) \quad (0.787) \quad (0.514)$$
$$H_t^2 = 0.075D_t - 0.336W_t + 0.346F_t + 0.868WN_t$$
$$(1.432) \quad (-6.437) \quad (0.450) \quad (1.370)$$

The first factor was composed of two significant variables (*D* and *W*), while the second one had only one significant variable (*W*). The non-parametric regression of *P* on the *N*´2 *H* matrix suggested that the linking function between the price and the first factor may be linear and that the second factor had a small effect on the conditional expectation of the price. Figure 2 illustrate the first factor. The SIR-NW algorithm captures the main features of the industry pricing behavior in the symmetric Cournot model with linear demand and a fixed number of firms with constant marginal cost, as described in equation (3).

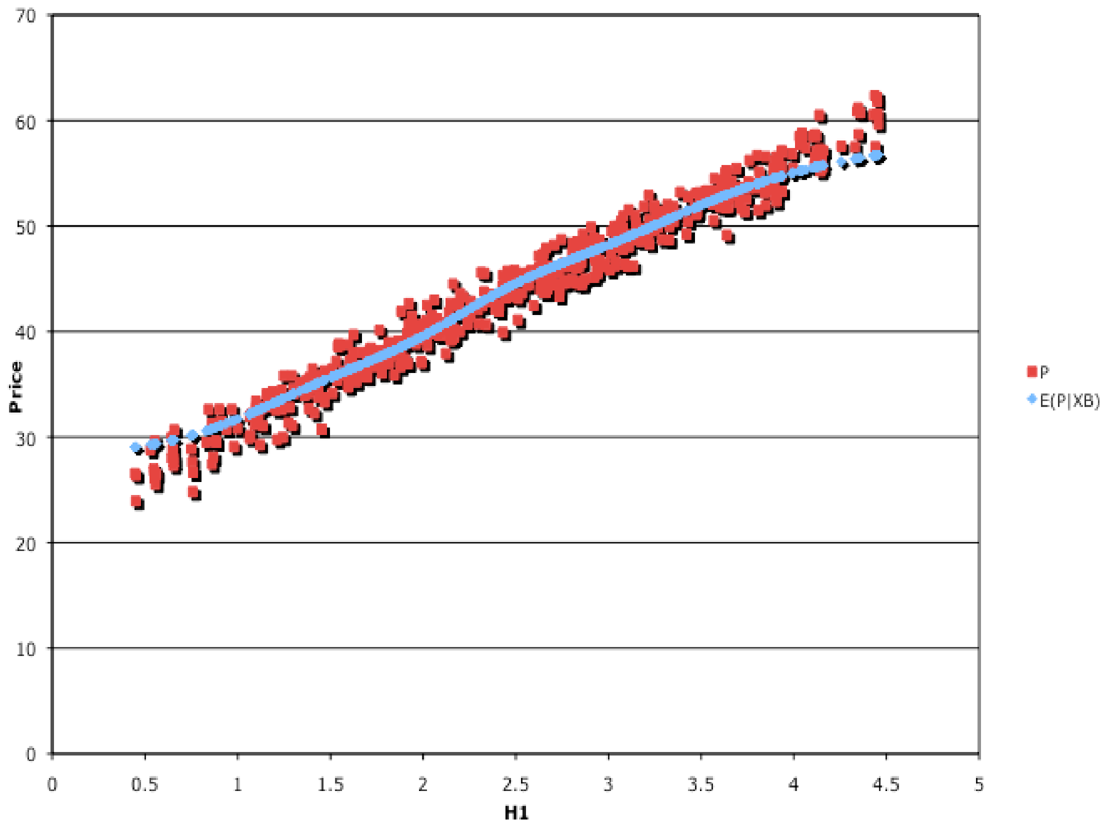

Figure 2. Non-parametric regression of price on the first SIR factor in a Cournot model with fixed number of firms.

In the Rotemberg-Saloner model, the collusive behavior of the firms is subject to an incentive constraint. The constraint creates a structural break in the data-generating process for the equilibrium quantity. As shown in equation (4), the industry produces the monopoly quantity if the constraint is not binding, or a larger volume if the constraint is binding. The change in the pricing depends on the state of nature described by the realizations of D and W. If D_t is large or/and W_t is small, firms have more incentive to deviate from the collusive agreement to take advantage of favorable market conditions. This behavior is particularly difficult to detect, because it requires that the estimation is able to separate the effect of a change in the shifters into its two components: the potential regime switch and the change in the optimal quantity in each regime. For example, an increase in the demand shifter increases the probability of having a binding constraint and increases the quantity produced in both regimes. The SIR-NW algorithm represents this change in the market regime with a kink in the estimated link function. The dashed line in Figure 3 illustrates the change in the slope in link function at the value of the first factor where the incentive constraint begins to bind.

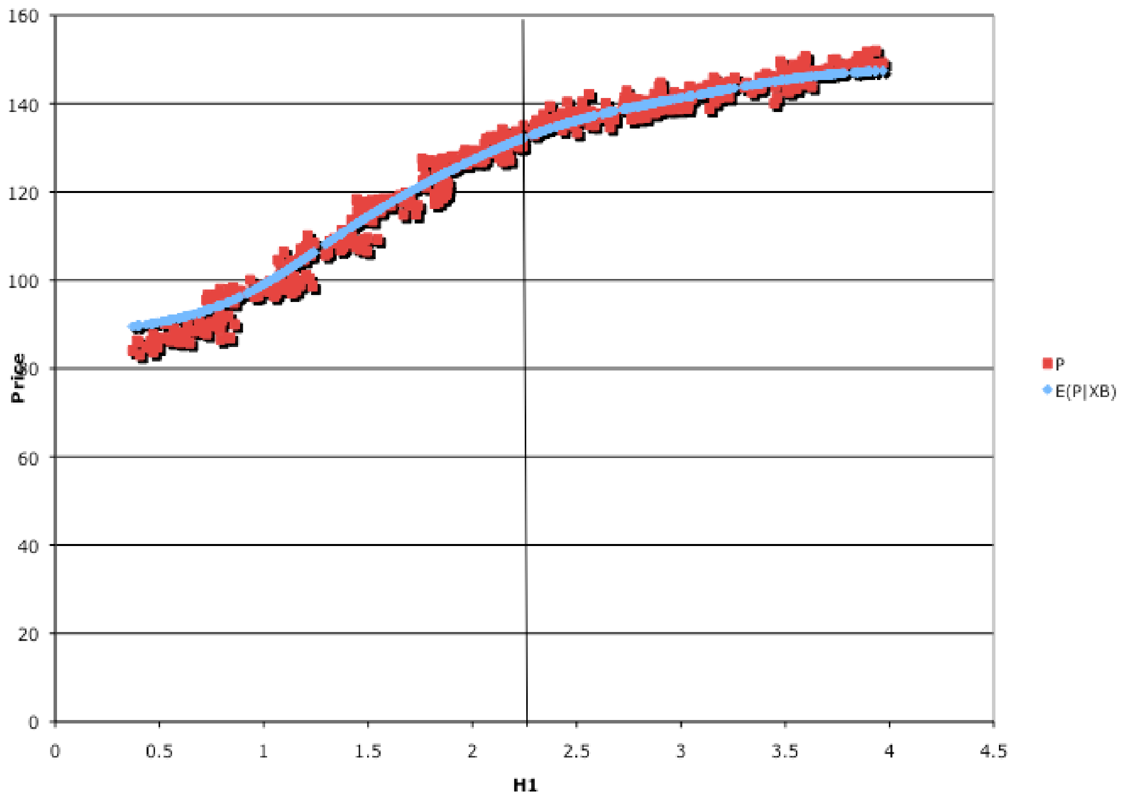

Figure 3. Non-parametric regression of price on the first SIR factor in a Rotemberg-Saloner model.

The analysis of the simulated data based on the three behavioral models shows that the SIR-NW algorithm was able to describe the specific features of each type of economic behavior. In particular, it was able to identify the significant variables determining price and changes in pricing behavior. Because this information is obtained without imposing assumptions on the data, it can guide the choice of the behavioral model used as the basis for estimation.

4 Conclusions

This essay presented a data-driven methodology for obtaining information about the industry's pricing behavior. The approach uses the SIR-NW algorithm, a two-step procedure which applies a sliced inverse regression to identify the significant factors affecting pricing behavior and then uses a non-parametric regression to estimate the corresponding link function. While it is not a replacement for a structural model of industry behavior, it provides information that can aid in choosing a behavioral model ex ante and assessing the specification bias ex post.

The SIR-NW algorithm is particularly useful when the econometrician has little a priori information regarding the industry behavior, because in this case the risk of misspecification in structural models is high. The exploratory approach allows the researcher to compensate for the lack of prior knowledge of the data-generating mechanism by obtaining information directly from the data.

5 References

Appelbaum, E. (1982). "The Estimation of the Degree of Oligopoly Power." Journal of Econometrics 19(2): 287-299.

Bresnahan, T. (1987). "Competition and Collusion in the American Automotive Industry: The 1995 Price War." The Journal of Industrial Economics 35(4): 457-482.

Bresnahan, T. (1989). Empirical Studies of Industries with Market Power. Handbook of Industrial Organization. R. Schmalensee and R. Willig, Elsevier. 2: 1011-1057.

Carter, C. and D. MacLaren (1997). "Price or Quantity Competition? Oligopolistic Structures in International Commodity Markets." Review of International Economics 5(3): 373-385.

Corts, K. (1999). "Conduct Parameters and the Measurement of Market Power." Journal of Econometrics 88: 227-250.

Cox, D. R. (1961). Tests of Separate Families of Hypotheses. Procedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley.

Dockner, E. (1992). "A Dynamic Theory of Conjectural Variation." The Journal of Industrial Economics 40(4): 377-395.

Dong, F., L. Marsh and K. W. Stiegert (2006). "State Trading Enterprises in a Differentiated Product Environment: The Case of Global Malting Barley Markets." American Journal of Agricultural Economics 88(1): 90-103.

Gasmi, F., J. Laffont and Q. Vuong (1992). "Econometric Analysis of Collusive Behavior in a Soft-Drink Market." Journal of Economics & Management Strategy 1(2): 277-311.

Karp, L. S. and J. S. Perloff (1989). "Dynamic Oligopoly in the Rice Export Market." The Review of Economics and Statistics 71(3): 462-479.Karp, L. S. and J. S. Perloff (1993). "Open-Loop and Feedback Models of Dynamic Oligopoly." International Journal of Industrial Organization 11: 369-389.

Katchova, A., I. Sheldon and M. Miranda (2005). "A Dynamic Model of Oligopoly and Oligopsony in the U.S. Potato-Processing Industry." Agribusiness 21(3): 409-428.

Makowski, L. (1987). "Are "Rational Conjectures" Rational?" The Journal of Industrial Economics XXXVI(1): 35-47.

Nevo, A. (2001). "Measuring Market Power in the Ready-to-Eat Cereal Industry." Econometrica 69(2): 307-342.

Pollak, R. and T. Wales (1991). "The Likelihood Dominance Criterion." Journal of Econometrics 47: 227-242.

Raper, K. C., A. H. Love and R. C. Shumway (2000). "Determining Market Power Exertion between Buyers and Sellers." Journal of Applied Econometrics 15(3): 225-252.

Rotemberg, J. and G. Saloner (1986). "A Supergame-Theoretic Model of Price Wars during Booms." American Economic Review 76(3): 390-407.

Sheldon, I. and R. Sperling (2003). "Estimating the Extent of Imperfect Competition in the Food Industry: What Have We Learned?" Journal of Agricultural Economics 54(1): 89-109.

Slade, M. (1995). "Empirical Games: the Oligopoly Case." Canadian Journal of Economics 28(2): 368-402.