COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 935

TESTING GAME-THEORETIC MODELS OF PRICE-FIXING BEHAVIOUR

Vassilis A. Hajivassiliou

January 1990

Testing Game–Theoretic Models of Price–Fixing Behaviour

by

Vassilis A. Hajivassiliou*

Cowles Foundation for Research in Economics
Yale University

December 1989

Abstract

This paper analyzes price-fixing by the Joint Executive Committee railroad cartel from 1880 to 1886 and develops tests of two game-theoretic models of tacit collusion. The first model, due to Abreu, Pearce and Stacchetti (1986), predicts that price will switch across regimes according to a Markov process. The second, by Rotemberg and Saloner (1986), postulates that price-wars are more likely in periods of high industry demand. Switching-regressions are used to model the firms' shifting between collusive and punishment behaviour. The main econometric novelty in the estimation procedures introduced in this paper is that misclassification probabilities are allowed to vary endogenously over time. The JEC data set is expanded to include measures of grain production to be shipped and availability of substitute transportation services. Our findings cast doubt on the applicability of the Rotemberg and Saloner model to the JEC railroad cartel, while they confirm the Markovian prediction of the Abreu et al. model.

* This research has been partly financed by NSF #SES 8813630. I would like to thank David Roth for expert research assistance and for helpful discussions. I have benefited from constructive comments by Jim Heckman, Lung-Fei Lee, Al Klevorick, Nancy Lutz, Ariel Pakes, David Pearce, Julio Rotemberg, Anne Royalty, Mike Whinston, and by participants at seminars at Harvard/M.I.T., Yale, Berkeley, Stanford, Virginia, and Munich.

TESTING GAME–THEORETIC MODELS OF PRICE–FIXING BEHAVIOUR

1. INTRODUCTION

This paper attempts to test two recent game–theoretic models of cartel behaviour. The first model, due to Abreu, Pearce, and Stacchetti (1986), predicts that price will switch across collusive/price–warfare regimes according to a Markov process. The second model of cartel behaviour I will test is due to Rotemberg and Saloner (1986) and predicts that the probability of a price-war is higher in periods of high industry demand. Maximum likelihood (ML) and simulation estimation methods that allow for measurement errors in switching-regression models are presented and applied to test these models using data on the Joint Executive Committee (JEC) railroad cartel for the period between 1880 and 1886.

Section 2 discusses the two game-theoretic models of tacit collusive behaviour, the Abreu, Pearce, and Stacchetti (1986) model and the Rotemberg and Saloner (1986) model, that will be tested using the techniques of this paper. The approach I develop here to test the basic Markovian prediction of the Abreu et al. (1986) model allows me to employ more information than do Berry and Briggs (1988), in the form of imperfect regime–classification information. Such information is available from several sources. The Rotemberg and Saloner (1986) model predicts that price-wars are more likely when industry demand is *high*. This seems counter to the conventional view of the classical industrial organization literature (see for example Stigler (1964)).

The JEC data for this study are discussed in Section 3. The data set, originally used by Porter (1983b) and Lee and Porter (1984), is expanded for the first time to include measures of grain production to be shipped and the availability of substitute transportation services. The construction of these indices is described in Appendix 1.

Section 4 analyzes *measurement errors in regime-classification information.* In this case the ML estimator that treats imperfect classifying information as perfect is inconsistent (Lee and Porter (1984)). Moreover, ML estimators that do not use regime–

classifying information are in general seriously inefficient (Goldfeld and Quandt (1975)). In all econometric applications of the switching-regression methodology, the misclassification problem is ubiquitous (see Lee and Porter (1984) for an analysis of cartel stability and Lee (1978) for an analysis of unionization). I examine appropriate estimators that incorporate imperfect classification information in the form of (multiple) indicator variables. A major difference between my procedure and the Lee and Porter (1984) analysis is that mine allows probabilities of misclassification *to vary* over the sample period. This feature is a priori expected to be crucial once exogenous classifying information is available, because one normally expects greater difficulties in accurately classifying a market when it is closer to a transition period.[1] The identification of the econometric model with varying misclassification probabilities is shown in Appendix 2.

Section 5 discusses estimation methods for switching models with imperfect classification information when switching occurs according to a Markov process. These methods are needed to test the predictions of the Abreu et al. game-theoretic model of firm behaviour.[2] I derive a recursion relation that makes evaluation of the likelihood function tractable. The details of this algorithm are described in Appendix 3.

A second type of measurement error, which is studied in Section 6, is of the classical variety: *errors in measuring explanatory variables.* Since the indices for industry demand and for extra-cartel competition constructed in this paper may contain serious measurement errors, it is important that estimation methods allow for this possibility. I show that simulation estimation methods (McFadden (1989), Pakes and Pollard (1989)) can handle the concomitant high-dimensional integrals that arise in non-linear errors-in-variables models such as the switching-regression models of this paper.

The estimation procedures are employed in Section 7 to analyze the price-fixing behaviour of the JEC railroad cartel. I find that the predictions of the Rotemberg and

---

[1] The methodology is in the spirit of the Tobit model of Nelson (1977) with a stochastically unobservable threshold.

[2] For a switching-regime model with the switching following a Markov process, see Cosslett and Lee (1985), whose results provide indirect support for the Abreu et al. theory.

Saloner model are not borne out by this data set, while the evidence favours the Markovian switching prediction of the Abreu et al. theory. Section 8 offers concluding remarks.

2. GAME–THEORETIC MODELS OF PRICE–FIXING BEHAVIOUR

Lee and Porter (1984) and Porter (1983b) used switching-regression methodology to test the game–theoretic models of Porter (1983a) and Green and Porter (1984). In this paper I will consider two other models of price-fixing, one by Rotemberg and Saloner (1986) and the second by Abreu et al. (1986). In Sections 4,5, and 6 I will introduce an econometric methodology to test the models while allowing for measurement errors.[3]

First, let me summarize the Rotemberg and Saloner (1986) model. Consider a symmetric n-firm, price-setting cartel facing stochastic demand. At each period the level of demand is a random variable independently and identically distributed (i.i.d.) over time. The firms learn the realized state of demand *before* making (simultaneously) their price-choices. When demand is high, there is an important temptation to undercut because by the i.i.d. assumption it is expected that demand will be lower next period. Hence, a punishment by the competitors would appear less severe than if high demand were believed likely to persist next period. As a result, Rotemberg and Saloner (1986) predict that in the presence of observable demand shocks, price wars (in the sense of lower amount of collusion) will mostly occur during industry booms.[4] This prediction is contrary to the conventional wisdom of the traditional industrial organization literature, which holds that it is generally more difficult to collude successfully during industry recessions when each firm is possibly preoccupied with its own survival.[5] Rotemberg and Saloner (1986),

---

[3] For an excellent review of game-theoretic models of tacit collusion see chapter 6 in Tirole (1988).

[4] These are not necessarily price-wars in the usual sense of periods of the maximal punishment of Bertrand (competitive) behaviour, because the price may actually be higher during booms than during recessions.

[5] On the other hand, Scherer (1980), p.222 discusses a case from the antibiotics industry where the discipline of the cartel was sorely tested by the arrival of a large order.

pp.395–396, describe how the basic version of their model can be modified to give behaviour fluctuating between periods of cooperation and non-cooperation, by further restricting the strategy space so that the oligopoly can choose only between the joint monopoly and the competitive prices. They label this version of their model as "intuitively more appealing," and show that the basic prediction of price-wars more likely to occur in demand booms is preserved.

The alternative game-theoretic model of collusion that is to be subjected to econometric tests is due to Abreu et al. (1986). In this model, the firms do not observe their competitors' quantities, but rather the market price whose distribution is determined by industry output and realized demand. Demand shocks are i.i.d. over time, but are not observed by the firms. In this basic supergame, the firms have concave objective functions and the distribution of the market price p_t conditional on aggregate output Q_t is assumed to have the property that a low price is more likely to have arisen from a high Q_t than from a low one. Abreu et al. (1986) are able to show that under these assumptions price-wars will result, with the behaviour of the firms being characterized by a trigger scheme, usually a "tail test."[6] During periods of successful collusion, firms will be producing q^+, earning payoff V^+. A trigger-price level p^+ will be determined such that an observation of a price lower than this will trigger a punishment phase, in which (higher) output q^- will be produced earning the firms a lower payoff V^-. This provides incentives for firms to restrict output. A second trigger p^- will determine whether a punishment phase will persist, or whether the industry will revert to successful collusion. Since a harsh punishment requires a high output, reversion to successful collusion will involve an "inverse tail test:" if a high price (greater than p^-) is observed, the game remains in the punishment phase.

[6] One of the key additional conditions required for the trigger scheme to be a simple "tail test" is the *monotone likelihood ratio property* (MLRP), defined as

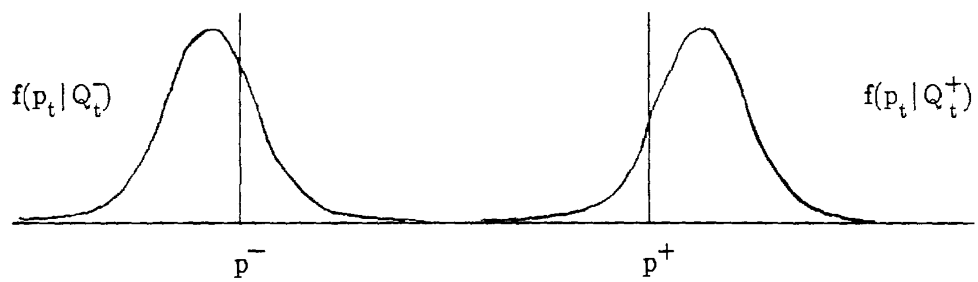
$$\frac{\partial}{\partial p_t}\left[\frac{\partial f/\partial Q_t}{f}\right] < 0 ,$$

where $f \equiv f(p_t \mid Q_t)$.

Conversely, successful collusion resumes when a price is observed lower than the p^- threshold; again, this provides incentives for high output.

Figure 1 gives two examples of the distribution of price given total output $f(p_t|Q_t)$, one when aggregate output is collusive (Q^+) and one when the industry is producing high output (Q^-); the trigger prices p^+ and p^- also appear in the figure. Let $I_t=1$ denote collusive output in period I_t and $I_t=0$ denote a punishment period.

Figure 1



The Abreu et al. model predicts industry behaviour switching between periods of successful collusion and punishment phases (of endogenous duration) according to a Markov process. If period t–1 was one of successful collusion, then with probability $\text{Prob}(p_t<p^+|p_{t-1}>p^+)$ a punishment phase will begin in period t. Given a period t–1 of punishment behaviour ($I_{t-1}=0$), the cartel will continue its high output/low price punishment with probability $\text{Prob}(p_t<p^-|p_{t-1}<p^+)$ and with probability $\text{Prob}(p_t>p^-|p_{t-1}<p^+)$ will resume successful collusion, if the punishment phase began in period t–1. If the industry has been in the punishment state for more than one period, these probabilities will be $\text{Prob}(p_t<p^-|p_{t-1}>p^-)$ and $\text{Prob}(p_t>p^-|p_{t-1}>p^-)$ respectively.[7] For another variant of

[7] Pursuing the implications of this theory further, a positive shift in the distribution of demand due to some exogenous factor would increase the persistence in the Markov process: Imagine a rightward shift of the $f(p_t|Q_t)$ family of distributions, which the firms do not perceive and hence continue employing the same p^+,p^- thresholds. It is straightforward to see that *both* $\text{Prob}(I_t=1|I_{t-1}=1) \equiv \text{Prob}(p_t>p^+|p_{t-1}>p^+)$ *and* $\text{Prob}(I_t=0|I_{t-1}=0) \equiv \text{Prob}(p_t>p^-|p_t<p^+)$ will *rise* under such a scenario.

such a model with similar qualitative implications, consider the price-secrecy model of Tirole (1988), section 6.7., which predicts price-wars as triggered by recessions. In this model, the firms who produce a differentiated product do not observe their competitors' prices, but try to infer them from their own demand. This assumption is in line with Stigler (1964). This model also has the prediction that the industry switches between collusive and punishment (in this case Bertrand) phases. Hence, in this theory, price-wars can be involuntary, in that they may be triggered by an unobservable negative demand shock and not necessarily by secret undercutting by a cartel member.

An econometric implementation of a simple model of a symmetric oligopoly that is able to nest in a simple way the Rotemberg and Saloner (1986) and the Abreu et al. (1986) game-theoretic models of price-fixing is the following. Suppose that an n-firm price-setting cartel switches between collusive and punishment (or non-collusive) behaviour, according to

(2.1) $$I_t = R_t(\Omega_t),$$

where Ω_t is the relevant information set available to the firms at time t and I_t takes the value 0 if punishment (or non-collusive) behaviour occurs in period t and the value 1 if collusion occurs then. Further suppose that the industry is characterized by a marginal cost function $MC_t=f(z_t,\epsilon_{ct})$ and a demand function $q_t=q(p_t,x_t,\epsilon_{dt})$, where z_t and x_t are vectors of exogenous variables. The shocks ϵ_c and ϵ_d may or may not be observed by the firms at the time of the decision, depending on the model under analysis, but are always unobservable by the econometrician. The functions $f(\cdot)$ and $q(\cdot)$ and the exogenous variables are known to the firms. In competitive periods price equals marginal cost, while in collusive periods marginal revenue equals marginal cost. Hence, in the latter periods, $p_t=f(z_t,\epsilon_{ct}) - q_t(p_t,x_t,\epsilon_{dt})/q_t'(p_t,x_t,\epsilon_{dt})$, where $q_t'(p_t,x_t,\epsilon_{dt}) \equiv \frac{\partial}{\partial p_t} \{q_t-(p_t-,x_t,\epsilon_{dt})\}$. The evolution of p_t can thus be summarized by

(2.2) $$p_t = f(z_t,\epsilon_{ct})-I_t \cdot q(p_t,x_t,\epsilon_{dt})/q'(p_t,x_t,\epsilon_{dt}) ,$$

where $I_t \equiv 1$(collusion is effective in period t).

For our econometric implementations I choose to work with the following parameterization of the demand function:[8]

$$\text{(2.3)} \qquad q_t = a.e^{-p_t/g(x_t,\epsilon_{dt})} .$$

This functional form is chosen so that price will be independent of x_t in competitive periods, but will vary positively with $g(x_t)$ in collusive ones. For tractability, assume further that $g(x_t,\epsilon_{dt})=e^{x_t\beta}+\epsilon_{dt}$, and that $MC_t=f(z_t)+\epsilon_{ct}$. It then follows that $q_t/q_t'=-g(x_t,e_{dt})$. I also assume that ϵ_c and ϵ_d are independent of one another, independent over time, and (conditionally) independent from x_t and z_t.[9] The price and (log) quantity equations for observation t in the two regimes can then be shown to be:

$$\begin{array}{lll} \text{(2.4.0)} & I=0 & p=f(z)+\epsilon_{p0} \\ & & \ln(q)=\ln(a)-f(z)/e^{x\beta}+\epsilon_{q0} \\ \text{(2.4.1)} & I=1 & p=f(z)+e^{x\beta}+\epsilon_{p1} \\ & & \ln(q)=\ln(a)-(1+f(z)/e^{x\beta})+\epsilon_{q1} . \end{array}$$

The game-theoretic models under study differ primarily in their implications about the switching rule (2.1). Let W_t denote all the relevant exogenous variables in the model. Abreu et al. (1986) predict that switching between regimes will evolve according to a Markov process. This theory can be parameterized by

---

[8] See Roth (1988) for a variant of this model that also allows parametrically different degrees of collusion. The model presented here possesses the property that output under collusion is 1/e times the competitive one.

[9] This will imply that $E(\epsilon_{p0}|X,Z)=E(\epsilon_{p1}|X,Z)=E(\epsilon_{q0}|X,Z)=E(\epsilon_{q1}|X,Z)=0$, and hence the main requirement for consistency of the ML procedures will be satisfied. This specification will also imply certain contemporaneous covariance structure between ϵ_{p0}, ϵ_{p1}, ϵ_{q0}, and ϵ_{q1} in (2.4). In this version of the paper, I neglect the particular form of this covariance structure; ignoring it will only affect the efficiency of the estimation procedures. I plan to exploit this issue in future refinements.

$$(2.5.a) \qquad I_t = 1 \quad \text{if} \quad W_t\gamma+\rho I_{t-1}+u_t \geq 0$$
$$= 0 \quad \text{otherwise.}$$

On the other hand, according to Rotemberg and Saloner (1986), the extent of collusion in some variants of their model, or the probability of a switch into a collusive regime in others, falls as the level of industry demand increases. A simple parameterization of this prediction is

$$(2.5.b) \qquad I_t = 1 \quad \text{if} \quad W_t\gamma+u_t \geq 0$$
$$= 0 \quad \text{otherwise,}$$

where W_t includes the demand variables x_t in the model, and the signs of the elements of γ corresponding to these variables are the opposite to the signs of the coefficients these variables have in the demand function.[10]

3. THE JEC DATA

The data analyzed in this study are 328 weekly observations on the operation of the Joint Executive Committee (JEC) railroad cartel, which primarily shipped grain from Chicago to the East Coast, over the period from week 1, 1880 to week 16, 1886. As documented by Ulen (1979), the cartel had extremely varied success in setting price and sharing the market. The effective price charged by each cartel member was not perfectly observable by its rivals because there is evidence that special shipping rates were sometimes secretly arranged with selected customers. The official price of shipping grain, labeled as series GR, is plotted in Figure 2. According to MacAvoy (1965), there were two critical periods of non-adherence: most of 1881 and most of 1884-1885. For detailed historical discussions of the events, see MacAvoy (1965), Ulen (1979), and Roth (1988).

[10] For example, a variable with a positive demand effect ($\beta>0$) should have a negative effect on the probability of successful collusion ($\gamma<0$).

The only exogenous information used in the Porter (1983b) and Lee and Porter (1984) studies was a dummy variable indicating whether the lakes were open for navigation. A plot of this variable appears in Figure 3a. This information is important because lake traffic was the leading substitute for shipping by railway. I here compile and use two additional pieces of exogenous information: an index of extra-cartel railroad competition and an index of total grain produced in the Midwest liable to be shipped to the East Coast. For details on the construction of these indices, see Appendix 1. The index of extra-cartel competition was based on the simple assumption that the strength of such competition was positively related to the number of railroads shipping grain to the East Coast that were operating outside the cartel. The index of total grain production in the midwest is a value-weighted annual total of the three largest grain crops produced in eight midwestern states, linearly interpolated to obtain weekly values for the index. The total Midwestern grain production index appears in Figure 3b.

In addition, various sources of information about regime classification are available. Ulen (1983) and MacAvoy (1965) constructed such indicators by relying on perceptions of the effectiveness of the JEC cartel as reported in contemporaneous weekly trade periodicals. See Figure 3b for the index of cartel-adherence compiled by Ulen (1979).[11] Porter (1983b) constructed a third index based on the predictions from his econometric model, according to the criterion of maximum probability. In Figure 2, Porter's (1983b) predictions of a price-war appear as series PWPorter. These regime classification indicators will be incorporated in the econometric implementation of the tests in Section 7, using the multiple dummy indicator models developed in the next sections.

Lack of good technological variables forced us to adopt the further simplifying assumption of constant (apart from stochastic shocks) marginal cost $MC_t=\alpha_0+\epsilon_{pt}$, which implies that $W_t=x_t$.[12] Four exogenous variables comprise x_t: a vector of ones, the dummy

[11] In Figure 2, this information also appears in the form of the series PWUlen, which gives the price-war periods according to this index.

[12] A simple alternative we plan to explore is that there are seasonal effects in the marginal cost function.

variable indicating whether or not the lakes were open, the index of extra-cartel railway competition, and the index of grain available to be shipped.

4. IMPERFECT CLASSIFICATION INFORMATION IN SWITCHING MODELS

Consider the general switching-regression model:

(4.1.a) $\quad y^*_{it} = h_i(X_i\delta_i)+\epsilon_{it} \qquad i = 0,1; \qquad t = 1,\ldots,T$

(4.1.b) $\quad y^*_{2t} = h_2(Z_t\gamma)+\epsilon_{2t}$

(4.1.c) $\quad y^*_{3t} = y^*_{2t}+\eta_t$.

y^*_{0t}, y^*_{1t}, y^*_{2t}, and y^*_{3t} are latent variables, unobservable by the econometrician. X_0, X_1, X_2, and Z are matrices of explanatory (exogenous) variables. ϵ_{0t}, ϵ_{1t}, ϵ_{2t}, η_t are normally distributed, i.i.d. over time, zero-mean disturbances. The functions $h_i(\cdot)$ are known to the econometrician up to the vectors of parameters δ_i and γ which will be estimated.

The econometrician observes the (endogenous) variable Y_t, which is generated as follows:

$$Y_t = y^*_{1t} \qquad \text{iff} \quad y^*_{2t} \geq 0$$

(4.2) $$\quad = y^*_{0t} \qquad \text{otherwise.}$$

In standard terminology, the two equations (4.1.a), i=0,1, are termed the "switched" equations and (4.1.b) the "switching" equation.

Let 1(A) be the indicator function, taking the value 1 iff logical condition A is true, 0 otherwise. Using this function to define the dummy variables $I_t \equiv 1(y^*_{2t} \geq 0)$ and $D_t \equiv 1(y^*_{3t} \geq 0)$, the econometrician observes D_t but not I_t. As long as $\sigma^2_\eta > 0$, D_t is an imperfect measurement of I_t. In this sense, η_t can be thought of as *coding error.*

In its general form without measurement errors, this model was used by Lee (1978) to study union/nonunion wage determination. A person's wages are denoted by y^*_{1t} if he/she is in a union, and by y^*_{0t} if non-union. The worker's propensity to be a union member is represented by y^*_{2t}. Fair and Jaffee (1972), *inter alia*, used the model to analyze markets in disequilibrium, by letting y^*_{1t} denote notional demand in period t, y^*_{0t} notional supply, and y^*_{2t} excess demand. As Lee and Porter (1984) explain, using inaccurate regime classification information in ML estimation leads to inconsistency. Moreover, Goldfeld and Quandt (1975) show that if perfect information is not used, ML estimation is seriously inefficient.[13]

Lee and Porter (1984) allowed for a *constant* probability of misclassification of observations into the two regimes because their only explanatory variable in the switching equation, Z, was a constant. Assuming a constant probability of misclassification once one has exogenous Z information is inappropriate, however, given that one expects the probability of misclassification to vary over time. In this paper, I model the probability of misclassification as a monotonic function of the (unobservable) degree of the propensity of the industry to lie in a particular regime. For example, in the disequilibrium version of the switching model, it seems plausible to assume that the probability of misclassification is smaller the larger the level of excess demand in the system. As I shall demonstrate, the coding error equation (4.1.c) incorporates this property into the model.

The contribution of an (independent) observation t to the likelihood function of the switching-regression model with coding error can then be derived as follows:

$$
(4.3)\qquad
\begin{array}{lllll}
\text{for } D_t=1: & (\, y^*_{3t}\geq 0\,) & \text{if } y^*_{2t}\geq 0,\ \eta_t\geq -y^*_{2t} & Y_t=y^*_{1t} & (\, I_t=1\,) \\
 & & \text{if } y^*_{2t}<0,\ \eta_t\geq -y^*_{2t} & Y_t=y^*_{0t} & (\, I_t=0\,) \\
 & & & & \\
\text{for } D_t=0: & (\, y^*_{3t}<0\,) & \text{if } y^*_{2t}<0,\ \eta_t< -y^*_{2t} & Y_t=y^*_{0t} & (\, I_t=0\,) \\
 & & \text{if } y^*_{2t}\geq 0,\ \eta_t< -y^*_{2t} & Y_t=y^*_{1t} & (\, I_t=1\,)\,.
\end{array}
$$

Let us use the notation $p_{d|it}\equiv \text{prob}(D_t=d|I_t=i)$, $p_{dit}\equiv \text{prob}(D_t=d,I_t=i)$, $p_{dt}=\text{prob}(D_t=d)$,

---

[13] Hajivassiliou (1987) combines these results to derive Hausman (1978) tests of accuracy of classification information.

π_{it}=prob(I_t=i), and f_{it}=pdf(y^*_{it}), where d and i take values 0 or 1. For simplicity assume that ϵ_{0t} and ϵ_{1t} are independent of ϵ_{2t} and η_t.[14] Dropping the t subscript for simplicity, this specification implies that the log-likelihood contribution is:

$$(4.4) \qquad \text{prob}(D,y|X) = D\cdot\ln(p_{1|1}f_1+p_{1|0}f_0)+(1-D)\cdot\ln(p_{0|1}f_1+p_{0|0}f_0)\ .$$

Note that the $p_{d|i}$'s involve bivariate integrals of the form

$$(4.5) \qquad p_{d|i} = \iint_{S_{DI}} f(\epsilon_2,\theta)d\epsilon_2 d\theta \Big/ \int_{S_I} f(\epsilon_2)d\epsilon_2\ ,$$

where $\theta\equiv\epsilon_2-\eta$, and the regions of integration (as described in (4.3)) are the sets:

$$S_{DI} = \{\ \epsilon_2 \underset{I}{\gtrless} -Z\gamma\ ,\ \eta \underset{D}{\gtrless} -(Z\gamma+\epsilon_2)\ \}\ \text{ and }\ S_I = \{\ \epsilon_2 \underset{I}{\gtrless} -Z\gamma\ \}\ ,$$

$$\text{where } \underset{I}{\gtrless} \equiv \{\ \geq \text{if } I=1,\ < \text{if } I=0\ \} \text{ and } \underset{D}{\gtrless} \equiv \{\ \geq \text{if } D=1,\ < \text{if } D=0\ \}.$$

The customary distributional assumption of normality is imposed.[15]

The coding error model with likelihood function defined in (4.3)-(4.5) possesses the desired property that the misclassification probability is highest at the borderline case when a regime switch appears most likely, and falls monotonically as the classifying exogenous information becomes stronger. To see this, first note that the probabilities of misclassification are:

$$(4.6) \qquad \begin{array}{ll} (D=1|I=0): & p_{1|0} = \text{prob}(\eta_t\geq-y^*_{2t}\ |\ y^*_{2t}<0\ ) \\ (D=0|I=1): & p_{0|1} = \text{prob}(\eta_t<-y^*_{2t}\ |\ y^*_{2t}\geq0\ )\ . \end{array}$$

Figure 4 presents probability plots for the misclassification case of D=1 and I=0, as a

[14] This assumption can be relaxed at the cost of further computational complexity.

[15] Specifically, I assume that $\begin{bmatrix}\epsilon_1\\ \epsilon_0\\ \epsilon_2\\ \theta\end{bmatrix} \sim N(\underset{\sim}{0}, \begin{bmatrix} \sigma_0^2 & \rho_{01}\sigma_0\sigma_1 & \rho_{02}\sigma_0 & 0 \\ & \sigma_1^2 & \rho_{12}\sigma_1 & 0 \\ & & 1 & \sigma_\eta \\ & & & 1+\sigma_\eta^2 \end{bmatrix})$.

It should be noted that ρ_{01} is unidentified and hence normalized to 0. Also for identification, σ_2 is normalized to 1.

function of the exogenous part of the switching equation, $Z\gamma$. Various values of the standard deviation of the coding error η are considered. As can be seen from Figure 4a, the conditional probabilities of misclassification, $p_{0|1}$ and $p_{1|0}$, are monotonic in $Z\gamma$ in the correct direction, rising when the signal $Z\gamma$ tends to suggest the wrong regime more strongly. For example, when the true state of the system is no collusion (I=0), higher values of $Z\gamma$ are further at odds with the truth, hence Prob(D=1|I=0) rises. As the standard deviation of the coding error η rises, the signal becomes less informative; in the limit, when $\sigma_\eta \to \infty$, the misclassification probabilities (Prob(D=d|I=i), $d \neq i$,) approach 0.5 . Hence, we confirm that the switching model with coding error introduced here possesses the desired property that the misclassification probability falls as the tendency to lie in a particular regime rises. In Figure 4b we see that the joint probability of misclassification p_{10} has a unique mode at the least informative value of the signal, $Z\gamma=0$, since in such a case it is most difficult to correctly classify the particular period.

An important caveat is that the coding-error switching-regression model allows only a limited degree of systematic misclassification. For example, despite the presence of the coding errors, note that the only change in the discrete part of the model is in the variance of the unobservable, which is of course unidentified. This is illustrated in Figure 4c. Hence, one can obtain consistent estimates for γ up to scale despite such misclassification.[16] This, of course, does not imply that the presence of the coding error is unimportant, because ML estimation of the complete discrete/continuous switching-regression model would still yield inconsistent results if the measurement errors were neglected.

Finally, suppose we have M multiple indicators $D_1,\ldots,D_M$ of regime classification. This is the non-linear analogue of the classic MIMIC model of Goldberger (1972). We then

[16] The importance of this restrictive feature of our measurement errors model will be investigated in future work.

obtain 2^{M+1} categories with respect to $D_1,...,D_M$, and I.[17] The likelihood contributions will in general involve (M+1)-fold integrals, which can be calculated by numerical methods for M up to 2 or 3. This modelling approach also has the desirable property that the misclassification probabilities vary over the sample period depending on the true probability of switching.[18]

[17] For the purposes of the empirical implementation in Section 7 with two imperfect classification indicators, I define $R \equiv Z\gamma$ and give the eight possibilities in that case:

D_1	D_2	I			
1	1	1	$\epsilon_2-\eta_1\leq R$	$\epsilon_2-\eta_2\leq R$	$\epsilon_2\leq R$
1	1	0	$\epsilon_2-\eta_1\leq R$	$\epsilon_2-\eta_2\leq R$	$\epsilon_2>R$
1	0	1	$\epsilon_2-\eta_1\leq R$	$\epsilon_2-\eta_2>R$	$\epsilon_2\leq R$
1	0	0	$\epsilon_2-\eta_1\leq R$	$\epsilon_2-\eta_2>R$	$\epsilon_2>R$
0	1	1	$\epsilon_2-\eta_1>R$	$\epsilon_2-\eta_2\leq R$	$\epsilon_2\leq R$
0	1	0	$\epsilon_2-\eta_1>R$	$\epsilon_2-\eta_2\leq R$	$\epsilon_2>R$
0	0	1	$\epsilon_2-\eta_1>R$	$\epsilon_2-\eta_2>R$	$\epsilon_2\leq R$
0	0	0	$\epsilon_2-\eta_1>R$	$\epsilon_2-\eta_2>R$	$\epsilon_2>R$.

[18] Under these normality assumptions, the bivariate integrals are calculated through an algorithm of Divgi (1979) in the empirical implementations below. In the case of two indicator variables in Section 7, the implied trivariate integrals are calculated through the method in Steck (1958).

5. A MARKOVIAN SWITCHING MODEL WITH IMPERFECT CLASSIFICATION

The models of the previous section exhibited a Bernoulli switching structure, conditional on the exogenous variables, which is characterized by the transition matrix

$$(5.1)\qquad \begin{array}{lll} & I_t=1 & I_t=0 \\ I_{t-1}=1 & \tau_t & 1-\tau_t \\ I_{t-1}=0 & \tau_t & 1-\tau_t\,. \end{array} \qquad \text{Bernoulli}$$

The transition probabilities τ's depend on time only through the exogenous variables, but not on the past state variable. In this section I introduce a model that allows the switching process to exhibit Markov dependence over time. This is necessary to test the main Markovian prediction of the game-theoretic model of Abreu et al. (1986). The Markov process has the transition structure

$$(5.2)\qquad \begin{array}{lll} & I_t=1 & I_t=0 \\ I_{t-1}=1 & \tau_{11t} & 1-\tau_{11t} \\ I_{t-1}=0 & \tau_{10t} & 1-\tau_{10t}\,, \end{array} \qquad \begin{array}{c}\text{Markov}\\ (\tau_{11}>\tau_{10})\end{array}$$

where $\tau_{ijt}=\text{Prob}(I_t=i\,|\,I_{t-1}=j)$. Specifically, to introduce a Markov structure of order 1, I modify the switching equation (4.1.b) so that the propensity to switch y^*_{2t} depends on the lagged state I_{t-1}, i.e.,

$$(5.3)\qquad \begin{aligned} I_t &= 1 \quad \text{if} \quad Z_t\gamma+\rho I_{t-1}+\epsilon_{2t}\geq 0 \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

With perfect classification information, this structure is straightforward to estimate since

$$(5.4)\qquad \begin{aligned} p(\underset{\sim}{y},\underset{\sim}{I}\,|\,X) &= p(y_T,I_T\,|\,I_{T-1},X_T)\cdot p(y_{T-1},I_{T-1}\,|\,I_{T-2},X_{T-1}) \\ &\quad \cdots p(y_2,I_2\,|\,I_1,X_2)\cdot p(y_1,I_1\,|\,I_0,X_1)\cdot p(I_0)\,.\text{[19]} \end{aligned}$$

The likelihood function for process (5.2–5.3), however, becomes extremely

[19] Note that how one treats $p(I_0)$ is not crucial, since this term has asymptotically vanishing influence. This is in contrast to the longitudinal data set case.

intractable in the presence of imperfect regime-classification information because it will require the evaluation of 2^T terms. The reason is as follows. We can readily show that

$$(5.5.a) \quad p(D_t,I_t|I_{t-1}) = I_t p(D_t,I_t=1|I_{t-1})+(1-I_t)p(D_t,I_t=0|I_{t-1})$$

$$(5.5.b) \quad p(y_t,D_t|I_{t-1}) = D_t[f_1 I_t p(1,1|I_{t-1})+f_0(1-I_t)p(1,0|I_{t-1})] + (1-D_t)[f_1 I_t p(0,1|I_{t-1})+f_0(1-I_t)p(0,0|I_{t-1})],$$

where I_t is determined by (5.3). But the econometrician only observes D_t , given by

$$(5.6) \qquad D_t = 1 \quad \text{if} \quad Z_t\gamma+\rho I_{t-1}+\epsilon_{2t}+\eta_t \geq 0$$
$$= 0 \quad \text{otherwise.}$$

Since I_{t-1} is unobserved by the econometrician for all t, the likelihood function is

$$(5.7) \qquad p(\underset{\sim}{y},\underset{\sim}{D}|X) = \sum_{I_T}\sum_{I_{T-1}}\ldots\sum_{I_2}\sum_{I_1}\sum_{I_0} p(y_T,D_t,I_T|I_{T-1})\cdots p(y_1,d_1,I_1|I_0).p(I_0) .$$

Due to the fact that each pair of consecutive terms involves I_{t-1} , the likelihood $p(\underset{\sim}{y},\underset{\sim}{D}|X)$ will in general require the evaluation of 2^T terms, a patently intractable task when T is of the order of 300, as in this paper. To solve this problem I show in Appendix 3 that, by extending ideas in Cosslett and Lee (1985) and Moran (1986), a recursion relation can be derived that makes evaluation of (5.7) feasible.

Note again that our approach here differs fundamentally from Lee and Porter (1984) and Cosslett and Lee (1985) in that the probability of misclassification is not constant but varies monotonically with the magnitude of $Z_t\gamma+\rho I_{t-1}$. A priori, this is a realistic feature. Given the dependence over time described in (5.2), one should expect the probability of misclassification to vary over time; it should be highest close to the boundary points when a switch occurs. These properties are exhibited by the conditional probability expressions above.[20]

[20] There is a cost, however, in terms of computational complexity because the conditional probability expressions $p(D_t,I_t|I_{t-1})$ now involve bivariate normal integrals (and in general (M+1)-fold integrals when M imperfect regime indicator variables are available).

6. CLASSICAL MEASUREMENT ERRORS

At least two explanatory variables that I shall use to test the game-theoretic models are expected *a priori* to be measured with potentially serious errors. The most suspect variables are the constructed index of grain production to be shipped and the constructed measure of the availability and strength of extra-cartel competition. Hence, in this section I investigate the effect of measurement errors in the explanatory variables of nonlinear models of the type estimated here. I will show that structural ML estimation of such models introduces high-dimensional integration problems, and that simulation estimation methods (McFadden (1989), Pakes and Pollard (1989)) avoid these difficulties.

To see how multiple integrals enter in models of the kind analyzed in this paper, consider the general limited dependent variable (LDV) model:

$$y_t^* = z_t^*\beta + \epsilon_t\,, \tag{6.1}$$

where the econometrician observes $y_t=\tau(y_t^*)$. $\tau(\cdot)$ is the function that maps the vector of underlying latent variables, y_t^*, into the vector of observable endogenous variables, y_t. For example, in the switching-regression model of Section 4, $y_t^*=(y_{0t}^*,y_{1t}^*,y_{2t}^*,y_{3t}^*)'$, $y_t=(Y_t,D_t)$, and the $\tau(\cdot)$ function is specified implicitly in equation (4.3). z_t^* is a k×1 vector of explanatory variables that are not directly observed. Instead we have the imperfect measurement x_t, given by

$$x_t = z_t^*\delta + v_t\,. \tag{6.2}$$

In terms of the observable limited dependent variable y_t, the model can be written as:

$$y_t = \tau(z_t^*\beta+\epsilon_t)\,. \tag{6.3}$$

Assuming that the measurement errors are of the classical form, it is plausible to postulate that $f(y|z^*,x)=f(y|z^*)$, which follows from mean-independence of the measurement error v_t and the true variables z_t^*.[21] Hence, using basic properties of conditional probability functions, the likelihood contribution conditional on the observable vector x is

$$(6.4) \qquad f(y|x) = \int f(y,z^*|x)dz^* = \int f(y,z^*,x)/f(x)dz^* = \int f(y|z^*,x)\cdot f(z^*|x)dz^* = \int f(y|z^*)\cdot f(z^*|x)dz^* .$$

Equation (6.4) illustrates that integration of order equal to the number of x variables will be needed to evaluate the likelihood in terms of the observable variables y and x. This is *in addition* to any integration required to calculate the density of y conditional on the unobservable z^* variables. These difficulties arise because the mismeasured regressors appear inside the non-linear function $\tau(\cdot)$. For example, in the switching-regression models of the previous sections, one must first calculate the likelihood assuming all explanatory variables are observed without error, $f(y|z^*)$. Then double numerical quadrature is required to evaluate $f(y|x)$, given that two of the z^*'s are observed imperfectly.[22] In the estimation section below, I offer both quadrature-based ML estimates with two mismeasured explanatory variables, as well as estimates by simulation estimation methods. This allows a comparative evaluation of the latter.[23]

---

[21] An alternative way of introducing measurement errors in the explanatory variables is through *mixture models*. This approach simultaneously estimates non-parametrically the (x,z^*) relation. Given the relatively small sample size for my complicated nonlinear model, I chose instead to parameterize explicitly the measurement errors to be of the classical type. For a general review of measurement error models, see Fuller (1987).

[22] A second difficulty that arises in nonlinear models with classical measurement errors in RHS variables is that the covariance matrix of the $f(y|z,x)$ distribution is a general $k\times k$ matrix Σ, even when ϵ_t has a scalar covariance matrix. This problem is explained in a previous version of this paper.

[23] For a general simulation estimation method for limited dependent variable models with panel data that is continuous in the parameter vector, see Hajivassiliou and McFadden (1989).

7. RESULTS

Using the econometric framework presented in Section 4, let me summarize the switching model of cartel behaviour I use to test the two game-theoretic models:

Non–Collusive Behaviour:

(7.1.0) $I = 0$ $\quad$ $p=f(z)+\epsilon_{p0}$

$$\ln(q)=\ln(a)-f(z)/e^{x\beta}+\epsilon_{q0}$$

Collusive Behaviour:

(7.1.1) $I = 1$ $\quad$ $p=f(z)+e^{x\beta}+\epsilon_{p1}$

$$\ln(q)=\ln(a)-(1+f(z)/e^{x\beta})+\epsilon_{q1}\,.$$

Switching Equation:

(7.1.2) $I = 1$ $\quad$ if $y^*_{2t}=W_t\gamma+\rho I_{t-1}+u_t\geq 0$

$= 0$ $\quad$ otherwise.

Coding Error Equation:

(7.1.3) $D = 1$ $\quad$ if $y^*_{3t}=y^*_{2t}+\eta_t\geq 0$

$= 0$ $\quad$ otherwise.

The estimation results appear in Tables 1-3. The basic model 1NL uses the Ulen (1979) classification of regimes and allows for neither a Markov structure nor measurement errors. Model 2NL also uses Ulen's classification but employs the appropriate methodology of Section 4 to model it as an imperfect scheme. Note that ML estimation of Model 2NL requires the evaluation of bivariate normal integrals.

The main conclusion that can be drawn from Table 1 is that the treatment of the Ulen classification as perfect or imperfect has a serious impact on the estimates. The effect is summarized by a strongly significant variance for the coding error η. Despite the fact that no coefficient estimate switches sign once imperfections in the regime information are allowed, the predicted regime classifications change substantially. According to the criterion of maximum probability, Model 2NL predicts 122 out of 328 periods to represent collusion, compared to 101 by Model 1NL.[24]

The importance of lake traffic as a substitute is confirmed as its coefficient is negative and statistically significant. The cost parameter is statistically better determined once imperfections in the regime indicators are admitted. Our two new exogenous variables are very significant both on the demand side (β coefficients) and in the switching equation (γ's). Demand for railroad shipping by the cartel is higher when the lakes are closed, extra-cartel competition is ineffective, and more total grain is available to be shipped. Moreover, contrary to the Rotemberg and Saloner (1986) predictions, the variables that have positive demand effects raise the probability that the cartel is colluding effectively. The probability of collusion rises with the level of grain available for shipping and falls with effective outside competition and the lakes being open. In his detailed study of this cartel, Ulen (1983) also believes that cartel adherence was strongly positively correlated with cyclical demand conditions; for example, demand upturns appeared to have been enough to extricate the cartel from its price-warfare 1884–85 phase.[25]

Table 2 shows that the construction of the exogenous variables is probably not inducing serious measurement errors. Allowing for normal errors in the extra-cartel competition index does not affect results substantially, either in terms of the coefficient estimates, or of the predicted regime classifications.[26] An interesting by-product of Table 2 is a comparative evaluation of the MSM simulation estimation method of McFadden (1989)

---

[24] A third version was also estimated, model 3NL, which combined two sources of regime classifying information using the multiple indicator models developed above. The second regime indicator I tried was the one constructed by Porter (1983b) which employs the predictions from his estimated model. The results from the two-indicator model were very similar to those from the one-indicator model 2NL and are not reported.

[25] The cross-country analysis of Suslow (1988) using hazard modelling provides independent confirmation of the finding that successful collusion is more likely to occur in demand booms.

[26] An issue that remains unanswered is whether the assumption of the classical measurement errors in the explanatory variables through the approach of Section 6 is too restrictive. As already mentioned, one way to test for this possibility is through nonparametric mixture models. No such attempts are made in this paper, because of the high data requirements of such estimation approaches.

and Pakes and Pollard (1989). The ML results by numerical quadrature are very close to the MSM estimates when 100 replications were used for MSM.

Table 3 presents the results of estimating the cartel model in the presence of a Markov structure in the switching equation. The findings in Table 3 lend strong support to the Abreu et al. (1986) model, since the coefficient ρ that allows for the Markov structure is very strongly significant. The other parameters move substantially compared to those of the (conditional) Bernoulli model, suggesting that coefficient estimates under the (apparently untenable) assumption of Bernoulli switching should not be trusted. Moreover, the models with a Markov structure predict more periods to lie in the collusive regime, compared to the corresponding models estimated with ρ set to 0.[27] Longer lag structures were tried, but most of the time dependence does not seem to extend over two weeks. Some evidence against the optimal one-period punishment story can be seen (the asymptotic t-statistic being 1.82), but a caveat to be borne in mind is that a week may not be the economically relevant decision-making interval for this cartel.[28]

8. CONCLUSION

This paper tested key predictions of game-theoretic models of collusive behaviour by developing switching-regression models that allow for imperfect regime classification information and misclassification probabilities that are not constrained to be constant over time but vary monotonically with the underlying propensity to switch regimes. Econometric models were derived that exhibit a Markovian switching structure. To make

[27] It is possible that the finding of a strongly significant Markov structure may be caused by residual serial correlation in the unobservables. Unfortunately, even in linear models it is very difficult in practice to differentiate, through the implied common-factor restrictions, the presence of lagged dependent variables as regressors from residual serial correlation. Moreover, given the non-linearity of the models of this paper, explicit allowance for serial correlation is not feasible with ML methods, because integration of order T=328 would be required.

[28] The evidence in Ulen (1979) suggests that all price changes occurred on a Monday. But important time—aggregation issues of course remain.

estimation of these models feasible, I developed a recursive relation for evaluating the likelihood expressions. Special methods also allowed for the possibility that constructed indices of extra-cartel competition and of total grain production in the Midwest contained substantial measurement errors.

The game-theoretic models of collusive behaviour that were tested using these methods were those of Rotemberg and Saloner (1986) and of Abreu et al. (1986). The results favour strongly the Abreu et al. prediction of Markovian switching behaviour between punishment and collusive regimes. The results cast doubt on the key prediction of the Rotemberg-Saloner model that the probability of switching into collusion falls as the level of industry demand rises.

Other game-theoretic models of oligopolistic behaviour and tacit collusion relax the key assumption made by both Abreu et al. and by Rotemberg and Saloner that demand shocks are i.i.d. over time, and allow instead for serially correlated demand shocks. Such models are due to Riordan (1985) and Haltiwanger and Harrington (1988). It would be interesting to test econometrically whether the two strong findings of this paper, namely existence of a Markov structure in the switching behaviour and price-warfare being more likely in recessions, would survive such a generalization. I leave this issue to future research.

Appendix 1

Constructed Data Series

More details on the construction of these two series can be found in Roth (1988).

1. Construction of Extra–Cartel Railroad Competition Index

The strength of extra-cartel competition as a threat to the JEC cartel is assumed directly related to the number of extra-cartel railroads shipping grain to the East Coast. MacAvoy (1965) documents the existence of extra-cartel competing railroads and specifies the exact periods when the JEC cartel responded to the existence of such firms in each case. According to this information, the first extra-cartel competitor was acknowledged by JEC in week 210 of the sample (January 4, 1884); a second extra-cartel railroad appeared as a competitor on August 15, 1884 which is week 242; finally, a third railroad firm withdrew from the JEC cartel following an unfavourable ruling by a JEC arbitrator, and started competing with the cartel in February 6, 1885 (week 267). I follow Roth (1988) and assume that the strength of extra-cartel competition varied with the square-root of the number of railroads operating outside the cartel.

2. Construction of Total Midwest Grain Production Index

Annual data was collected on the largest grain crops (corn, wheat, and oats) from eight Midwestern states and weighted according to the average U.S. price for each grain over the period, to generate an annual value index of midwestern grain output. Considerations of a lag between harvest and shipping suggested assigning the annual grain production value to January 1 of the following year. Finally, simple linear interpolation was used to construct weekly values for this index.

Appendix 2

Identification of Measurement Error Switching Model

In this Appendix I show how all the parameters of the switching-regression model with coding error (4.1.a)–(4.1.c) are econometrically identified, subject to the normalization that $\sigma_2=1$. Recall the definitions

(A2.1.a) $\quad p_{di} \equiv \text{Prob}(D=d, I=i)$

(A2.1.b) $\quad p_{d|i} \equiv \text{Prob}(D=d \mid I=i)$

(A2.1.c) $\quad p_d \equiv \text{Prob}(D=d)$

(A2.1.d) $\quad \pi_i \equiv \text{Prob}(I=i)$.

Under the normality assumptions imposed and the normalization $\sigma_2=1$

$$\text{(A2.2)} \qquad p_d = D\cdot\text{Prob}(D=1) + (1-D)\cdot\text{Prob}(D=0) = D\cdot\Phi(Z\gamma/\sqrt{1+\sigma_\eta^2}) + (1-D)\cdot(1-\Phi(Z\gamma/\sqrt{1+\sigma_\eta^2}))$$

$$\text{(A2.3)} \qquad \pi_i = I\cdot\text{Prob}(I=1) + (1-I)\cdot\text{Prob}(I=0) = I\cdot\Phi(Z\gamma) + (1-I)\cdot(1-\Phi(Z\gamma))$$

Using only the imperfect classification indicator D, from the marginal likelihood (A2.2) we can estimate the expression $\gamma/\sqrt{1+\sigma_\eta^2}$ consistently.

Now consider the marginal likelihood for the observed endogenous variable y, neglecting any classification information, i.e., consider the marginal likelihood

$$\text{(A2.4)} \qquad f(y) = f_1\cdot\pi_1 + f_0\cdot\pi_0 = f_1\cdot\Phi(Z\gamma) + f_0\cdot(1-\Phi(Z\gamma)) .$$

From this, we can obtain consistent estimates for the parameters β_1, β_0, σ_1, σ_0, and γ, provided either the functions $h_0(\cdot)$ and $h_1(\cdot)$ are not the same, or further restrictions on β_1 and β_0 are imposed.

Finally, consider the conditional likelihood

$$\text{(A2.5)} \qquad f(y \mid D=1) = f_1\cdot\frac{p_{11}}{p_1} + f_0\cdot\frac{p_{10}}{p_1} ,$$

which uses separately the observations classified by the (imperfect) indicator D to be in collusion. We see immediately that in such a case the expressions p_{11}/p_1 and p_{10}/p_1 are consistently estimable. But

$$p_{11}/p_1 = p_{1|1} \cdot \pi_1/p_1 = p_{1|1} \cdot \Phi(Z\gamma)/\Phi(Z\gamma/\sqrt{1+\sigma_\eta^2}) \;; \tag{A2.6}$$

hence, $p_{1|1}$ can also be identified. The identification of the remaining term $p_{0|0}$ follows from exactly analogous arguments.

Appendix 3

A Recursion Algorithm

for the Markovian Switching–Regression Model with Coding Error

The aim is to facilitate the evaluation of the likelihood function of Section 5, which is given by:

$$\text{(A3.1)} \qquad p(\underset{\sim}{y},\underset{\sim}{D}\,|\,X) = \sum_{I_T}\sum_{I_{T-1}}\cdots\sum_{I_2}\sum_{I_1}\sum_{I_0} p(y_T,D_t,I_T\,|\,I_{T-1})\cdots p(y_1,d_1,I_1\,|\,I_0)\cdot p(I_0)\ .$$

The difficulty in evaluating (A3.1) directly is that each pair of consecutive terms involves I_{t-1}; hence, each likelihood evaluation will require calculating 2^T terms, which is a computationally prohibitive task.

The following arguments generalize ideas in Cosslett and Lee (1985) and Moran (1986) and show how (A3.1) can be evaluated recursively through T matrix multiplications. Define the set of available endogenous information at time t by S_t, i.e., $S_t \equiv (y_1,D_1,y_2,D_2,\ldots,y_t,D_t)$. Further define $Q_t(I_t)\equiv p(S_t,I_t)$. Since we can always write

$$\text{(A3.2)} \qquad Q_t(I_t) = p(S_{t-1},y_t,D_t,I_t) = \sum_{I_{t-1}} p(S_{t-1},I_{t-1},y_t,D_t,I_t)\ ,$$

it follows that

$$\text{(A3.3)} \qquad Q_t(I_t) = \sum_{I_{t-1}} p(y_t,D_t,I_t\,|\,I_{t-1},S_{t-1})\cdot p(I_{t-1},S_{t-1}) = \sum_{I_{t-1}} p(y_t,D_t,I_t\,|\,I_{t-1})\cdot Q_{t-1}(I_{t-1}),$$

where we have used the Markov structure $p(y_t,D_t,I_t\,|\,I_{t-1},S_{t-1})=p(y_t,D_t,I_t\,|\,I_{t-1})$ and the definition $Q_{t-1}(I_{t-1})\equiv p(I_{t-1},S_{t-1})$. But calculation of (A3.3) only requires information up to t, as the following matrix equation shows:

(A3.4)
$$\begin{bmatrix} Q_t(0) \\ \\ Q_t(1) \end{bmatrix} = \begin{bmatrix} p(y_t,D_t,I_t{=}0\,|\,I_{t-1}{=}0) & p(y_t,D_t,I_t{=}0\,|\,I_{t-1}{=}1) \\ \\ p(y_t,D_t,I_t{=}1\,|\,I_{t-1}{=}0) & p(y_t,D_t,I_t{=}1\,|\,I_{t-1}{=}1) \end{bmatrix} \begin{bmatrix} Q_{t-1}(0) \\ \\ Q_{t-1}(1) \end{bmatrix}$$

or,

$$\underset{\sim}{Q}_t = \underset{\sim}{M}_t \cdot \underset{\sim}{Q}_{t-1} \, .$$

The likelihood (A3.1) can thus be calculated recursively from (A3.4) and

(A3.1')
$$p(\underset{\sim}{y},\underset{\sim}{D}\,|\,X) = \sum_{I_T} Q_T(I_T) = Q_T(0) + Q_T(1).$$

TABLE 1

NO LAGGED STRUCTURE

(Asymptotic t-statistics in parentheses)
Regime Classification Variable Used: Ulen (1979)

Variable	Parameter	Model 1NL No lags, no measurement errors	Model 2NL No lags, 1 Imperfect Regime Indic.
	σ_0^2	6.13 (2.59)	5.63 (2.88)
	σ_1^2	0.17 (2.01)	7.23 (2.11)
	σ_η^2	——	1.63 (3.59)
<u>Marginal Cost Equation</u>:			
	α_0	16.63 (0.29)	8.21 (2.40)
<u>Demand Equation</u>:			
	ln(a)	10.96 (0.13)	11.96 (0.68)
	β_0	2.78 (0.11)	3.92 (2.27)
Lakes open dummy	β_1	–0.40 (–10.72)	–0.91 (–12.81)
Midwestern Grain Output	β_2	0.86 (11.12)	0.921 (9.812)
Extra JEC Competition	β_3	–0.44 (–15.53)	–0.34 (–11.43)
<u>Switching Equation</u>:			
	γ_0	–0.52 (0.32)	–0.72 (–1.95)
Lakes open dummy	γ_1	–0.76 (–2.77)	–0.93 (–4.83)
Midwestern Grain Output	γ_2	5.01 (4.95)	3.78 (7.23)
Extra JEC Competition	γ_3	–2.40 (–5.36)	–3.71 (–7.87)
Loglikelihood		–444.414	–437.342
Number of Periods Predicted			
in Collusive Regime		101.	122.
in Competitive Regime		227.	206.

TABLE 2

ERRORS IN EXPLANATORY VARIABLES

(Asymptotic t-statistics in parentheses)
Regime Classification Variable Used: Ulen (1979)

Variable	Parameter	Model 4NL*	Model 5NL
		Normal Errors in Extra Cartel Competition Variable, MSM	Normal Errors in Extra Cartel Competition Variable ML Quadrature
	σ_0^2	5.23 (3.66)	5.29 (3.06)
	σ_1^2	0.21 (2.72)	0.36 (2.89)
Marginal Cost Equation:			
	α_0	14.63 (0.72)	14.01 (0.63)
Demand Equation:			
	ln(a)	9.23 (0.27)	9.88 (0.42)
	β_0	2.97 (0.28)	3.21 (0.33)
Lakes open dummy	β_1	–0.77 (–11.25)	–0.88 (–12.38)
Midwestern Grain Output	β_2	-1.27 (9.21)	0.71 (8.99)
Extra JEC Competition	β_3	–0.22 (–11.83)	–0.23 (–10.35)
Switching Equation:			
	γ_0	–0.76 (–0.99)	–0.82 (1.22)
Lakes open dummy	γ_1	0.53 (–3.67)	–0.59 (–3.17)
Midwestern Grain Output	γ_2	4.83 (3.67)	4.61 (4.81)
Extra JEC Competition	γ_3	–1.83 (–7.48)	–2.01 (–7.12)
Loglikelihood		——	–453.231
Number of Periods Predicted			
in Collusive Regime		104.	105.
in Competitive Regime		224.	223.

*100 replications used in the simulation estimation

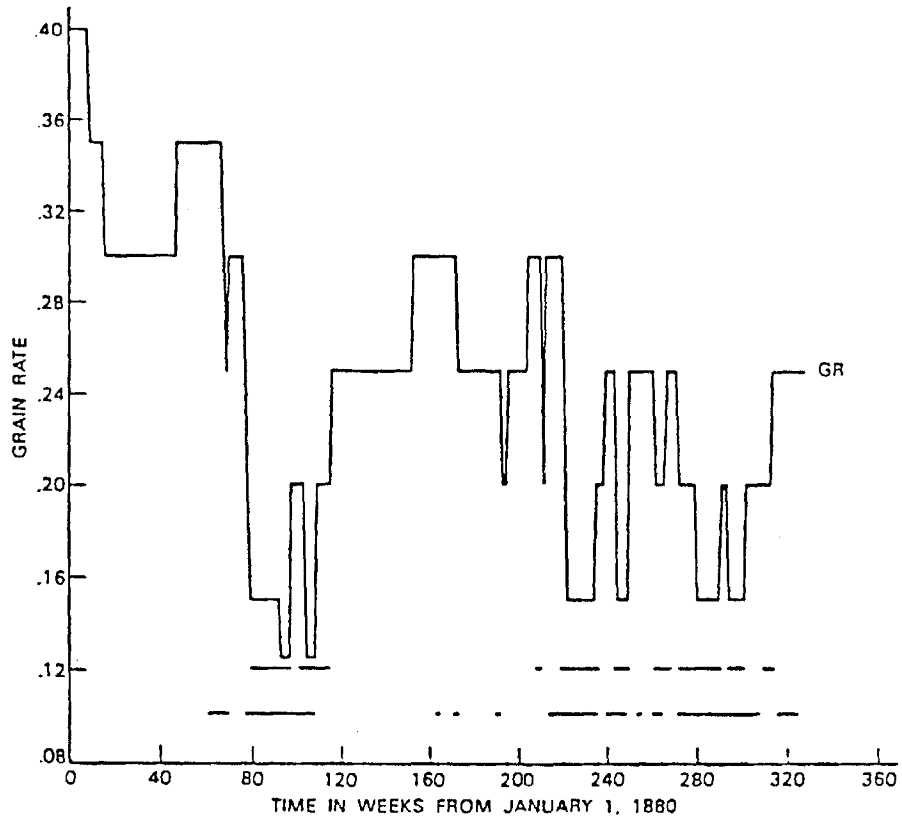
TABLE 3

MARKOV SWITCHING STRUCTURE

(Asymptotic t-statistics in parentheses)
Regime Classification Variable Used: Ulen (1979)

Variable	Parameter	Model 6L No errors, Markov structure	Model 7L No errors, Markov structure with 2 lags	Model 8L 1 imperfect dummy, Markov structure
	σ_0^2	6.97 (2.18)	6.78 (2.38)	6.32 (2.42)
	σ_1^2	0.78 (3.72)	0.81 (3.82)	0.64 (3.47)
Marginal Cost Equation:				
	α_0	12.63 (0.96)	12.52 (0.95)	10.43 (2.77)
Demand Equation:				
	ln(a)	7.93 (1.23)	7.83 (1.35)	7.42 (1.08)
	β_0	4.22 (0.74)	4.17 (0.81)	4.12 (0.72)
Lakes open dummy	β_1	–0.23 (–12.78)	–0.28 (–11.81)	–0.35 (–8.42)
Midwestern Output	β_2	0.72 (7.24)	0.77 (7.38)	0.81 (8.15)
Extra JEC Competition	β_3	–0.33 (–12.35)	–0.41 (–11.92)	–0.32 (–9.93)
Switching Equation:				
	γ_0	–0.50 (0.78)	–0.51 (0.82)	–0.47 (0.58)
Lakes open dummy	γ_1	–0.66 (–3.23)	–0.75 (–3.27)	–0.49 (–3.59)
Midwestern Grain Output	γ_2	3.37 (2.88)	3.42 (2.94)	3.75 (2.73)
Extra JEC Competition	γ_3	–2.93 (–6.72)	–2.74 (–6.84)	–3.02 (–6.32)
Regime Lagged Once	ρ_1	2.78 (6.23)	2.56 (5.72)	2.58 (6.52)
Regime Lagged Twice	ρ_2	——	0.25 (1.82)	——
Loglikelihood		–423.773	–422.661	–418.524
Number of Periods Predicted				
in Collusive Regime		113.	114.	129.
in Competitive Regime		215.	214.	199.

Figure 2



* Figure reproduced from Porter (1983b), p.311

Legend

GR	= Official price set by the JEC cartel for shipping grain
PWUlen	= Price-war classification according to Ulen (1979)
PWPorter	= Price-war classification according to Porter (1983b)

Figure 3

Fig 3a

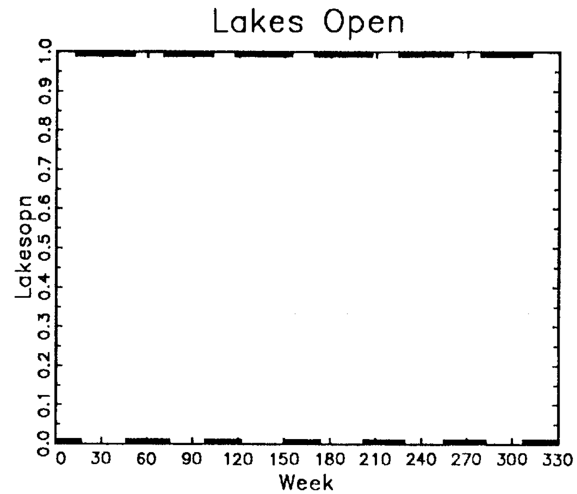


Fig 3b

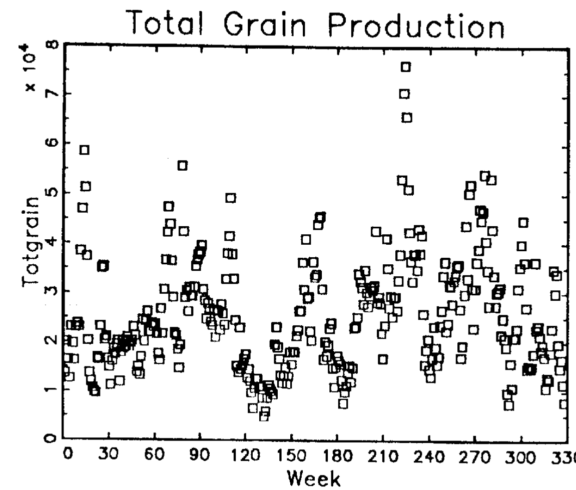


Fig 3c

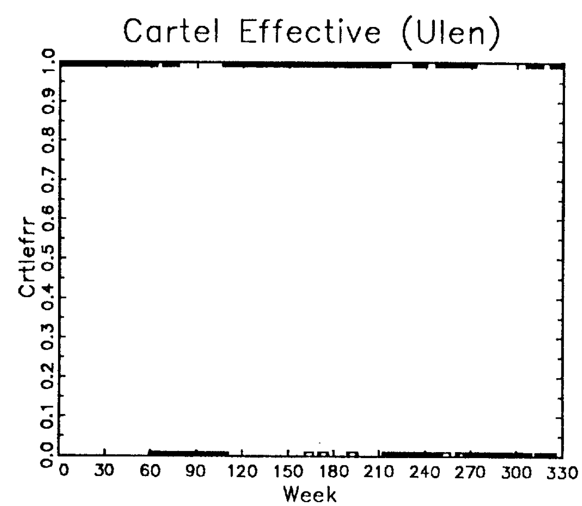


Figure 4

Fig 4a

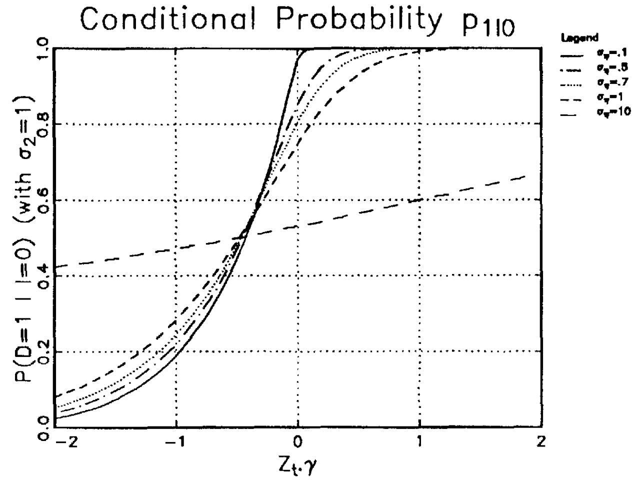


Fig 4b

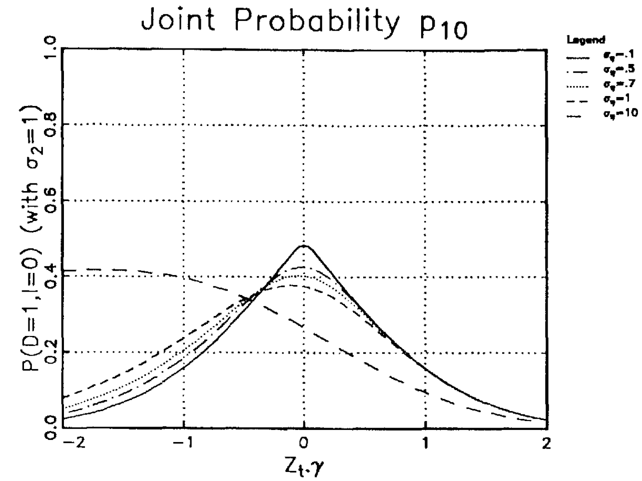
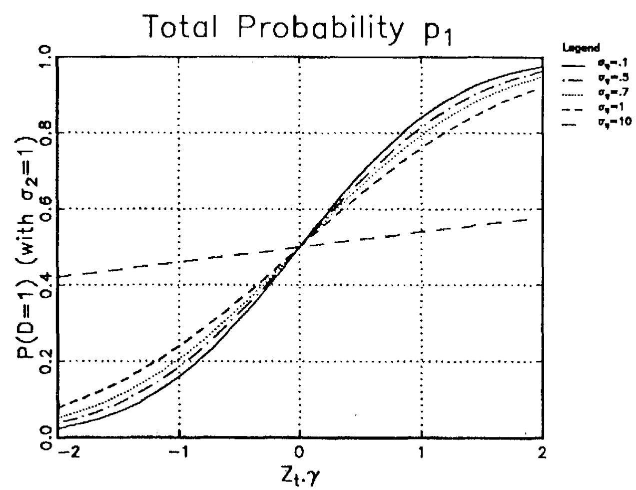


Fig 4c



REFERENCES

Abreu, D., D. Pearce, and E. Stacchetti (1986): "Optimal Cartel Equilibria with Imperfect Monitoring," Journal of Economic Theory 39, 251-269.

Berry, S. and H. Briggs (1988): "A Non-Parametric Test of a First Order Markov Process for Regimes in a Non-Cooperatively Collusive Industry," mimeo, University of Wisconsin, Madison.

Cosslett, S. and L.-F. Lee (1985): "Serial Correlation in Latent Discrete Variable Models," Journal of Econometrics 27, 79-97.

Divgi, D. (1979): "Calculation of Univariate and Bivariate Normal Probability Functions," Annals of Statistics, 903-910.

Fair, R. and D. Jaffee (1972): "Methods of Estimation for Markets in Disequilibrium," Econometrica 40.

Fuller, W. (1987): Measurement Error Models. Wiley: New York.

Goldberger, A. (1972): "Maximum Likelihood Estimation of Regressions Containing Unobservable Variables," International Economic Review 13, 1-15.

Goldfeld, S. and R. Quandt (1975): "Estimation of a Disequilibrium Model and the Value of Information," Journal of Econometrics 3, 325–348.

Green, E. and R. Porter (1984): "Non-Cooperative Collusion under Imperfect Price Information," Econometrica 52, 87-100.

Hajivassiliou, V.A. (1987): "Two Misspecification Tests for the Simple Switching Regression Model," Economics Letters 22, 343-348.

Hajivassiliou, V. and D. McFadden (1989): "The Method of Simulated Scores for the Estimation of LDV Models: With An Application to the Analysis of External Debt Crises," Cowles Foundation, Yale University, mimeo.

Haltiwanger, J. and J. Harrington, Jr. (1988): "The Impact of Cyclical Demand Movements on Collusive Behavior," University of Maryland working paper.

Hausman, J. (1978): "Specification Tests in Econometrics," Econometrica 46

Lee, L.-F. (1978): "Unionism and Wages: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables," International Economic Review 19, 415-433.

Lee, L.-F. and R. Porter (1984): "Switching Regression Models with Imperfect Sample Separation Information - with an Application on Cartel Stability," Econometrica 52, 391-418.

MacAvoy, P. (1965): The Economic Effects of Regulation. Cambridge: M.I.T. Press.

McFadden, D. (1989): "A Method of Estimating Multinomial Probits without Calculating Multiple Integrals," Econometrica 57, 995-1026.

Moran, (1986): "Orthant Probabilities and Gaussian Markov Processes," Journal of Applied Probability, 413–417.

Nelson, F. (1977): "Censored Regression Models with Unobservable Stochastic Thresholds," Journal of Econometrics 6, 309-328.

Pakes, A. and D. Pollard (1989): "The Asymptotics of Simulation Estimators," Econometrica 57, 1027-1058.

Porter, R. (1983a): "Optimal Carter Trigger Price Stability," Journal of Economic Theory 29, 313-338.

Porter, R. (1983b): "A Study of Cartel Stability: The Joint Executive Committee, 1880-1886," Bell Journal of Economics 14, 301-314.

Riordan, M. (1985): "Imperfect Information and Dynamic Conjectural Variations," Rand Journal of Economics 16, 41–50.

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

Roth, D. (1988): "The Joint Executive Committee Railroad Cartel of 1880-1886: A Reinterpretation and Respecification of Recent Empirical Work," Yale University, mimeo.

Scherer, F.M. (1980): Industrial Market Structure and Economic Performance. Boston: Houghton-Mifflin.

Steck, G. P. (1958): "A Table for Computing Trivariate Normal Probabilities," Annals of Mathematical Statistics 29, 780-800.

Stigler, G. (1964): "A Theory of Oligopoly," Journal of Political Economy 72, 44–61.

Suslow, V. (1988): "Stability in International Cartels: An Empirical Survey," University of Michigan mimeo.

Tirole, J. (1988): The Theory of Industrial Organization. Cambridge: MIT Press.

Ulen, T. (1979): "Cartels and Regulation: Railroad Collusion and the Origins of the Interstate Commerce Commission," unpublished Ph.D. dissertation, Stanford University.

Ulen, T. (1983): "Railroad Cartels before 1887: The Effectiveness of Private Enforcement of Collusion," Research in Economic History 8, 125–144.