Endogenous Switching Regressions with Imperfect Regime Classification Information: Theory and Applications

Vassilis Hajivassiliou
Department of Economics and FMG, LSE

CRETE2024 Conference, Milos, 12 July 2024

A Single Regime Indicator

Consider the general switching-regression model:

$$y_{it}^{*} = h_i(X_i\delta_i) + \epsilon_{it} \qquad i = 0, 1; \qquad t = 1, ..., T \tag{1.1.a}$$

$$y_{2t}^{*} = h_2(Z_t\zeta) + \epsilon_{2t} \tag{1.1.b}$$

$$y_{3t}^{*} = y_{2t}^{*} + \eta_t. \tag{1.1.c}$$

A Single Regime Indicator

Latent Variables: y^*_{0t}, y^*_{1t}, y^*_{2t}, and y^*_{3t} , unobservable by the econometrician;

Matrices of explanatory (exogenous) variables: X_0, X_1, X_2, and Z ; and

Multivariate normally distributed, $i.i.d.$ over time, with zero-mean disturbances: $(\epsilon_{0t}, \epsilon_{1t}, \epsilon_{2t}, \eta_t)'$

Functions $h_i(\cdot)$, $i = 0, 1, 2$, are known to the econometrician up to the vectors of parameters δ_i and ζ, which will be estimated.

A Single Regime Indicator

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

$$Y_t = \begin{cases} y^*_{1t}, & \text{iff } y^*_{2t} \geq 0 \\ y^*_{0t}, & \text{otherwise.} \end{cases} \tag{1.2}$$

In standard terminology, the two equations (1.1.a), $i = 0, 1$, are termed the "switched" equations and $(1.1.b)$ the "switching" equation. Using the indicator function $\mathbf{1}(\cdot)$, we define the dummy variables $I_t \equiv \mathbf{1}\left(y^*_{2t} \geq 0\right)$ and $D_t \equiv \mathbf{1}\left(y^*_{3t} \geq 0\right)$.
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 is *coding error*.

Econometric Novelties

- Endogenous Switching Regressions econometric modelling
- Novelty 1: with imperfect regime classification information.
- Novelty 2: misclassification probabilities allowed to vary endogenously over time.
- Standard MLE infeasible because likelihood contributions are sum of 2^T terms.
- Novelty 3: Algorithm developed that allows exact likelihood calculation through simply T matrix multiplications (each of a 2×2 matrix times a 2×1 vector.)

Econometric Novelties

- Endogenous Switching Regressions econometric modelling
- Novelty 1: with imperfect regime classification information.
- Novelty 2: misclassification probabilities allowed to vary endogenously over time.
- Standard MLE infeasible because likelihood contributions are sum of 2^T terms.
- Novelty 3: Algorithm developed that allows exact likelihood calculation through simply T matrix multiplications (each of a 2×2 matrix times a 2×1 vector.)

Econometric Novelties

- Endogenous Switching Regressions econometric modelling
- Novelty 1: with imperfect regime classification information.
- Novelty 2: misclassification probabilities allowed to vary endogenously over time.
- Standard MLE infeasible because likelihood contributions are sum of 2^T terms.
- Novelty 3: Algorithm developed that allows exact likelihood calculation through simply T matrix multiplications (each of a 2×2 matrix times a 2×1 vector.)

Econometric Novelties

- Endogenous Switching Regressions econometric modelling
- Novelty 1: with imperfect regime classification information.
- Novelty 2: misclassification probabilities allowed to vary endogenously over time.
- Standard MLE infeasible because likelihood contributions are sum of 2^T terms.
- Novelty 3: Algorithm developed that allows exact likelihood calculation through simply T matrix multiplications (each of a 2×2 matrix times a 2×1 vector.)

Econometric Novelties

- Endogenous Switching Regressions econometric modelling
- Novelty 1: with imperfect regime classification information.
- Novelty 2: misclassification probabilities allowed to vary endogenously over time.
- Standard MLE infeasible because likelihood contributions are sum of 2^T terms.
- Novelty 3: Algorithm developed that allows exact likelihood calculation through simply T matrix multiplications (each of a 2×2 matrix times a 2×1 vector.)

General Applicability

Methods shown widely applicable:

- Hamilton's work on Markov-Switching models;
- External financing problems faced by firms in Corporate Finance (Hajivassiliou and Savignac (2024): different investment functions and different innovation functions depending on whether or not firm faces binding finance constraints);
- Borrowing constraints faced by households impacting on their intertemporal consumption decisions (Hajivassiliou and Ioannides (2010): different intertemporal consumption functions for households depending on whether or not they face binding liquidity borrowing constraint);

General Applicability

Methods shown widely applicable:

- Hamilton's work on Markov-Switching models;
- External financing problems faced by firms in Corporate Finance (Hajivassiliou and Savignac (2024): different investment functions and different innovation functions depending on whether or not firm faces binding finance constraints);
- Borrowing constraints faced by households impacting on their intertemporal consumption decisions (Hajivassiliou and Ioannides (2010): different intertemporal consumption functions for households depending on whether or not they face binding liquidity borrowing constraint);

General Applicability

Methods shown widely applicable:

- Hamilton's work on Markov-Switching models;
- External financing problems faced by firms in Corporate Finance (Hajivassiliou and Savignac (2024): different investment functions and different innovation functions depending on whether or not firm faces binding finance constraints);
- Borrowing constraints faced by households impacting on their intertemporal consumption decisions (Hajivassiliou and Ioannides (2010): different intertemporal consumption functions for households depending on whether or not they face binding liquidity borrowing constraint);

General Applicability

- Aggregation over sectors in different levels of disequilibrium because of sluggish price adjustments (Hajivassiliou (2023));
- and
- Game-theoretic models of price collusion in Industrial Organization (this paper).

Draw contrasts and similarities to Regression Discontinuity Designs

General Applicability

- Aggregation over sectors in different levels of disequilibrium because of sluggish price adjustments (Hajivassiliou (2023));
- and
- Game-theoretic models of price collusion in Industrial Organization (this paper).

Draw contrasts and similarities to Regression Discontinuity Designs

General Applicability

- Aggregation over sectors in different levels of disequilibrium because of sluggish price adjustments (Hajivassiliou (2023));
- and
- Game-theoretic models of price collusion in Industrial Organization (this paper).

Draw contrasts and similarities to Regression Discontinuity Designs

Critical Advantages of the Switching Regressions Framework

The following critical advantages of the switching regressions framework need to be emphasized and more widely disseminated:

- Agents may **transition** endogenously between the alternative (switching) specifications depending on economic circumstances.
- The alternative switching specifications may **co-exist** at a given point in time across different agents, sectors of the economy, etc.
- In a panel data context, these two features may **both** apply.

Critical Advantages of the Switching Regressions Framework

The following critical advantages of the switching regressions framework need to be emphasized and more widely disseminated:

- Agents may **transition** endogenously between the alternative (switching) specifications depending on economic circumstances.
- The alternative switching specifications may **co-exist** at a given point in time across different agents, sectors of the economy, etc.
- In a panel data context, these two features may **both** apply.

Critical Advantages of the Switching Regressions Framework

The following critical advantages of the switching regressions framework need to be emphasized and more widely disseminated:

- Agents may **transition** endogenously between the alternative (switching) specifications depending on economic circumstances.
- The alternative switching specifications may **co-exist** at a given point in time across different agents, sectors of the economy, etc.
- In a panel data context, these two features may **both** apply.

Collusion Application:

Application here:

- Analyze price fixing by the Joint Executive Committee railroad cartel from 1880 to 1886.
- Develop tests of two prototypical game-theoretic models of tacit collusion.
- Model 1: Abreu, Pearce and Stacchetti (1986) "price will switch across regimes according to a Markov process"
- Model 2: Rotemberg and Saloner (1986) "price wars more likely in periods of high industry demand"

Collusion Application:

Application here:

- Analyze price fixing by the Joint Executive Committee railroad cartel from 1880 to 1886.
- Develop tests of two prototypical game-theoretic models of tacit collusion.
- Model 1: Abreu, Pearce and Stacchetti (1986) "price will switch across regimes according to a Markov process"
- Model 2: Rotemberg and Saloner (1986) "price wars more likely in periods of high industry demand"

Collusion Application:

Application here:

- Analyze price fixing by the Joint Executive Committee railroad cartel from 1880 to 1886.
- Develop tests of two prototypical game-theoretic models of tacit collusion.
- Model 1: Abreu, Pearce and Stacchetti (1986) "price will switch across regimes according to a Markov process"
- Model 2: Rotemberg and Saloner (1986) "price wars more likely in periods of high industry demand"

Collusion Application:

Application here:

- Analyze price fixing by the Joint Executive Committee railroad cartel from 1880 to 1886.
- Develop tests of two prototypical game-theoretic models of tacit collusion.
- Model 1: Abreu, Pearce and Stacchetti (1986) "price will switch across regimes according to a Markov process"
- Model 2: Rotemberg and Saloner (1986) "price wars more likely in periods of high industry demand"

Key Findings:

Key findings:

- Doubts as to applicability of Rotemberg and Saloner model to the JEC railroad cartel
- Confirmation of Markovian prediction of the Abreu et al. model

Key Findings:

Key findings:

- Doubts as to applicability of Rotemberg and Saloner model to the JEC railroad cartel
- Confirmation of Markovian prediction of the Abreu et al. model

Misclassfication of Regimes

- Literature: Lee and Porter (1984) allowed for *constant* probability of misclassification.
- Here: Misclassification probability endogenous.

Misclassfication of Regimes

- Literature: Lee and Porter (1984) allowed for *constant* probability of misclassification.
- Here: Misclassification probability endogenous.

Literature of A Single Regime Indicator

In its general form without measurement errors in regime classification, the switching-regression model was used by Lee (1978) to study union/nonunion wage determination.
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. Hajivassiliou (1987) combines these results to derive Hausman (1978) tests of accuracy of classification information.

Fixed Probability of Misclassification

Lee and Porter (1984) allowed for a *constant* probability that observations were misclassified into the two regimes; their only explanatory variable in the switching equation, Z, was a constant. **inappropriate if one expects the probability of misclassification to vary over time**, and especially so if one has exogenous information represented by Z_t, which, as theory suggests, should affect switching.

Endogenous Misclassification Probabilities

I model the misclassification probability as a monotonic function of the (unobservable) propensity of the industry to be in a particular regime measured by the latent variable y^*_{2t}. For example, in the disequilibrium version of the switching model (Fair and Jaffee (1972), it seems plausible to assume that the probability of misclassification is smaller the larger the level of excess demand in the system. I demonstrate shortly that the coding error equation $(1.1.c)$ incorporates this property into the model.

Endogenous Misclassification Probabilities

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

$$\begin{array}{llllll} \textit{for } D_t = 1: & (y^*_{3t} \geq 0) & \text{if } y^*_{2t} \geq 0, \eta_t \geq -y^*_{2t} & Y_t = y^*_{1t} & (I_t = 1) \\ & & \text{if } y^*_{2t} < 0, \eta_t \geq -y^*_{2t} & Y_t = y^*_{0t} & (I_t = 0) \\ \\ \textit{for } D_t = 0: & (y^*_{3t} < 0) & \text{if } y^*_{2t} < 0, \eta_t < -y^*_{2t} & Y_t = y^*_{0t} & (I_t = 0) \\ & & \text{if } y^*_{2t} \geq 0, \eta_t < -y^*_{2t} & Y_t = y^*_{1t} & (I_t = 1). \end{array} \tag{1.3}$$

Endogenous Misclassification Probabilities

Let us use the notation $p_{d|it} \equiv prob(D_t = d | I_t = i)$, $p_{dit} \equiv prob(D_t = d, I_t = i)$, $p_{dt} = prob(D_t = d)$, $\pi_{it} = prob(I_t = i)$, 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. This assumption can be relaxed at the cost of further computational complexity.

Dropping the t subscript for simplicity, this specification implies that the log-likelihood contribution is:

$$prob(D, y|X) = D \cdot \ln(p_{1|1} f_1 + p_{1|0} f_0) + (1 - D) \cdot \ln(p_{0|1} f_1 + p_{0|0} f_0). \tag{1.4}$$

Endogenous Misclassification Probabilities

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

$$p_{d|i} = \int \int_{S_{DI}} f(\epsilon_2, \theta) d\epsilon_2 d\theta \; / \int_{S_I} f(\epsilon_2) d\epsilon_2, \qquad (1.5)$$

where $\theta \equiv \epsilon_2 - \eta$, and the regions of integration are described in the paper and summarized here:

Endogenous Misclassification Probabilities

$S_{DI} = \{\epsilon_2 \underset{I}{\overset{>}{<}} -Z\zeta, \eta \underset{D}{\overset{>}{<}} -(Z\zeta + \epsilon_2)\}$ and $S_I = \{\epsilon_2 \underset{I}{\overset{>}{<}} -Z\zeta\}$, where $\underset{I}{\overset{>}{<}} \equiv \{\geq \text{ if } I = 1, < \text{ if } I = 0\}$ and $\underset{D}{\overset{>}{<}} \equiv \{\geq \text{ if } D = 1, < \text{ if } D = 0\}$.

Endogenous Misclassification Probabilities

The customary distributional assumption of normality is imposed. Semi- and non-parametric alternatives are currently explored with my PhD student Felix Iglehaut and with Camarena (2024 unpublished PhD dissertation at Birkbeck, UoL.

Endogenous Misclassification Probabilities

It should be noted that in the econometric models of this paper, there are two errors for the collusive regime, $\epsilon_1 = (\epsilon_{1p}, \epsilon_{1q})'$, and two for the punishment regime, $\epsilon_0 = (\epsilon_{0p}, \epsilon_{0q})'$. As discussed already in footnote 10, we will be neglecting the correlations between price and quantity errors for a given regime, $\rho(\epsilon_{rp}, \epsilon_{rq})$, $r = 1, 0$ (i.e., $\rho_{01} = 0$). Moreover, the correlations of errors across regimes cannot be identified, hence $\rho(\epsilon_{0i}, \epsilon_{1j})$ will be set to 0, $i, j = p, q$. Finally, also for identification, σ_2 is normalized to 1.

Endogenous Misclassification Probabilities

The coding error model with the likelihood function defined by using (1.3)–(1.5) 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 exogenous classifying information becomes stronger. To see this, first note that the probabilities of misclassification are:

$$\begin{aligned} (D=1|I=0): \quad & p_{1|0} = prob(\eta_t \geq -y^*_{2t}|y^*_{2t} < 0) \\ (D=0|I=1): \quad & p_{0|1} = prob(\eta_t < -y^*_{2t}|y^*_{2t} \geq 0). \end{aligned} \tag{1.6}$$

Endogenous Misclassification Probabilities

Figures 1-3 present probability plots for the misclassification case of $D = 1$ and $I = 0$ as a function of the exogenous part of the switching equation, $Z\zeta$. The corresponding plots for the case with $D = 0$ and $I = 1$ are exact mirror images with respect to $Z\zeta = 0$ of those in Figures 1-3 and are not given separately.
Various values of the standard deviation of the coding error η are considered.

Endogenous Misclassification Probabilities

As can be seen from Figure 1, the conditional probability of misclassification, $p_{1|0}$, is monotonic in $Z\zeta$ in the desired direction, rising when the signal $Z\zeta$ 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 $Z\zeta$ are further at odds with the truth, hence $\text{Prob}(D = 1|I = 0)$ 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.

Endogenous Misclassification Probabilities

In Figure 2 we see that the joint probability of misclassification p_{10} has a unique mode at the least informative value of the signal, $Z\zeta = 0$, 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, the only change in the discrete part of the model, $(1.1.b)$, is in the variance of the latent variable y^*_{2t}, which is, of course, unidentified.

Endogenous Misclassification Probabilities

This is illustrated in Figure 3. Hence, one can obtain consistent estimates for ζ up to scale despite such misclassification. The importance of this restrictive feature of my measurement errors model will be investigated in future work.

This, however, 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.

Multiple Regime Indicators

Finally, suppose we have M multiple indicators $D_1, \ldots, D_M$ of regime classification. This is the nonlinear analogue of the classic MIMIC model of Goldberger (1972). We then obtain 2^{M+1} categories with respect to $D_1, \ldots, D_M$, and I. For the purposes of the empirical implementation in Section 4 with two imperfect classification indicators, I define $R \equiv Z\zeta$ and give the eight possibilities in the paper.

Multiple Regime Indicators

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, like the coding-error model with a single indicator, (1.1)–(1.2), also has the desirable property that the misclassification probabilities vary over the sample period depending on the true probability of switching. Under the normality assumptions discussed above, in the empirical implementations below the bivariate integrals are calculated through an algorithm of Divgi (1979). In the case of two indicator variables in Section 4, the implied trivariate integrals are calculated through the method in Steck (1958). The higher dimension integrals implied by more indicators can be accommodated by simulation estimation methods. See Hajivassiliou (1993) for a discussion.

Markovian Switching Model with Imperfect Classification

Because of the *i.i.d.* assumptions on the error vector $(\epsilon_{0t}, \epsilon_{1t}, \epsilon_{2t}, \eta_t)'$, the models of the previous section exhibit a Bernoulli switching structure, conditional on the exogenous variables. This is characterized by a transition matrix that specifies probabilities for the transitions $I_{t-1} \longrightarrow I_t$.

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

A Markovian Switching Model with Imperfect Classification

When I_t is a Markov process, then the implied transition structure has the desired properties whereby the transition probabilities

$$\tau_{ijt} = Prob(I_t = i | I_{t-1} = j)$$

depend on the past state variable, as well as on the exogenous variables.

A Markovian Switching Model with Imperfect Classification

One expects positive serial persistence, in the sense of $\tau_{11t} > \tau_{10t}$. Specifically, to introduce a Markov structure of order 1, I modify the switching equation (1.1.b) so that the propensity to switch, y^*_{2t}, depends on the lagged state I_{t-1}, i.e.,

$$I_t = \begin{cases} 1 & \text{if } Z_t\zeta + \rho I_{t-1} + \epsilon_{2t} \geq 0 \\ 0 & \textit{otherwise}. \end{cases} \tag{5.3}$$

A Markovian Switching Model with Imperfect Classification

With perfect classification information, this structure is straightforward to estimate since[1]

$$p[\mathbf{Y}, \mathbf{I}, I_0 | \mathbf{X}] \equiv p[Y_1, \cdots, Y_T, I_1, \cdots, I_T, I_0 | X_1, \cdots, X_T]$$

$$= p[Y_T, I_T | I_{T-1}, X_T] \cdot p[Y_{T-1}, I_{T-1} | I_{T-2}, X_{T-1}] \cdots p[Y_2, I_2 | I_1, X_2] \cdot p[Y_1 \tag{5.4}$$

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

A Markovian Switching Model with Imperfect Classification

The likelihood function for the Markov process introduced in (5.3), however, becomes extremely intractable in the presence of imperfect regime-classification information because it then requires the evaluation of 2^T terms. The reason is as follows. We can readily show that

$$p[D_t, I_t | I_{t-1}] =$$

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

$$p[Y_t, D_t | I_{t-1}] =$$

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

where I_t is determined by (5.3).

A Markovian Switching Model with Imperfect Classification

But the econometrician only observes D_t, given by

$$D_t = \begin{cases} 1 & if Z_t\zeta + \rho I_{t-1} + \epsilon_{2t} + \eta_t \geq 0 \\ 0 & otherwise. \end{cases} \tag{5.6}$$

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

$$p[\mathbf{Y}, \mathbf{D}|\mathbf{X}] = \sum_{I_T}\sum_{I_{T-1}}\cdots\sum_{I_2}\sum_{I_1}\sum_{I_0} p[Y_T, D_t, I_T|I_{T-1}]\cdots p[Y_1, d_1, I_1|I_0]\cdot p[I \tag{5.7}$$

A Markovian Switching Model with Imperfect Classification

Because each pair of consecutive terms involves I_{t-1} , the likelihood $p[\mathbf{Y}, \mathbf{D}|\mathbf{X}]$ 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.

I solve this problem in Appendix 3 by extending ideas in Cosslett and Lee (1985) and Moran (1986), by deriving a recursion relation that makes evaluation of (5.7) feasible.

A Markovian Switching Model with Imperfect Classification

Note again that the approach here differs fundamentally from that of 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 $Z_t\zeta + \rho I_{t-1}$. 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. 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).

Models of Symmetric oligopoly

Rotemberg and Saloner (1986) and the Abreu et al. (1986)

Models of Symmetric oligopoly

$I_t = 0$ competitive (price-war) price equals marginal cost
$I_t = 1$ collusive period. marginal revenue equals marginal cost

Models of Symmetric oligopoly

Non-Collusive Behaviour:

$$q_t = ae^{-p_t/g(x_t,\epsilon_{dt})}. \tag{2.3}$$

$$I_t = 0 \begin{cases} p_t = MC(z_t) + \epsilon_{p0t} \\ \ln q_t = \ln a - MC(z_t)/e^{x_t\beta} + \epsilon_{q0t} \end{cases} \tag{2.4.0}$$

Collusive Behaviour:

$$I_t = 1 \begin{cases} p_t = MC(z_t) + e^{x_t\beta} + \epsilon_{p1t} \\ \ln q_t = \ln a - (1 + MC(z_t)/e^{x_t\beta}) + \epsilon_{q1t}. \end{cases} \tag{2.4.1}$$

Models of Symmetric oligopoly

Rotemberg and Saloner (1986) collusion or the probability of a switch into a collusive regime falls as the level of industry demand increases.

$$I_t = \begin{cases} 1 & \text{if } W_t\gamma + u_t \geq 0 \\ 0 & otherwise. \end{cases} \tag{2.5.a}$$

opposite sign that the coefficient of that variable has in the demand function ($\beta_j\gamma_j < 0$). For example, a variable with a positive demand effect ($\beta > 0$) should have a negative effect on the probability of successful collusion ($\gamma < 0$).

Models of Symmetric oligopoly

Abreu et al. (1986) switching between regimes will evolve according to a Markov process. Parameterized by

$$I_t = \begin{cases} 1 & if W_t\gamma + \rho I_{t-1} + u_t \geq 0 \\ 0 & otherwise, \end{cases} \qquad (2.5.b)$$

The JEC Data

328 weekly observations (1880-1886) Joint Executive Committee (JEC) railroad cartel

The only exogenous information used in the Porter (1983b) and Lee and Porter (1984) studies was whether or not the lakes were open for navigation. LakesOpen.

The JEC Data

I compiled and used two additional pieces of exogenous information: an index of extra-cartel railroad competition and an index of total grain produced in the Midwest that might be shipped to the East Coast. For details on the construction of these indices, see Appendix 1. The extra-cartel competition, which is plotted in Figure 1, is based on the simple assumption that the strength of such competition was positively related to the number of railroads that were shipping grain to the East Coast but 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 2

The JEC Data

Various other 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.

The JEC Data

Lack of good technological variables forces me to adopt the further simplifying assumption that, apart from stochastic shocks, marginal cost is constant, or $MC_t = \alpha_0 + \epsilon_{ct}$. This implies we have only demand variables to treat as exogenous, i.e., $W_t = x_t$.[2] Four such exogenous variables comprise x_t: a vector of ones, the dummy 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. *Ceteris paribus*, one expects higher demand for the shipping services of the cartel in periods when the lakes were closed to navigation, when competition from railways outside the cartel was not vigorous, and when there was a big grain crop in the Midwest.

[2] A simple alternative I plan to explore is that there are seasonal effects in

Results

I use the econometric framework of the switching model of cartel behaviour, presented in Section 1, to test the two game-theoretic models under the additional assumption that marginal cost does not depend on any exogenous information and is constant apart from a random error, i.e., $MC = \alpha_0 + \epsilon_{ct}$:

Results

Non-Collusive Behaviour:

$$I_t = 0 \begin{cases} p_t = \alpha_0 + \epsilon_{p0t} \\ \ln(q_t) = \ln(a) - \alpha_0 / e^{x_t \beta} + \epsilon_{q0t}. \end{cases} \tag{7.1.0}$$

Collusive Behaviour:

$$I_t = 1 \begin{cases} p_t = \alpha_0 + e^{x_t \beta} + \epsilon_{p1t} \\ \ln(q_t) = \ln(a) - (1 + \alpha_0 / e^{x_t \beta}) + \epsilon_{q1t}. \end{cases} \tag{7.1.1}$$

Switching Equation:

$$I_t = \begin{cases} 1 & if y_{2t}^* = W_t \gamma + \rho I_{t-1} + u_t \geq 0 \\ 0 & otherwise. \end{cases} \tag{7.1.2}$$

Coding Error Equation:

Results

In this specification, the parameter α_0 denotes the deterministic part of marginal cost; ln a and β are demand function parameters (see equation 2.3); and γ denotes the effect of the variables W_t on the likelihood of successful collusion in period t. If a variable x_j appears with a positive coefficient β_j, this means that variable x_j has a positive demand impact, *ceteris paribus*. Similarly, a variable w_j which makes collusion more likely should have a positive γ coefficient.

Conclusions

In this paper I have tested key predictions of two prototypical game-theoretic models of collusive behaviour by developing switching-regression models. The models allow for imperfect regime classification information and for misclassification probabilities that need not 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 estimate these models, 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 total grain production in the Midwest contained substantial measurement errors.

Conclusions

The prototypical game-theoretic models of collusive behaviour that were tested using these methods were those of Rotemberg and Saloner (1986) and 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.

Conclusions

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. These models, due to Riordan (1985) and Haltiwanger and Harrington (1988), allow instead for serially correlated demand shocks. 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 the greater likelihood of price warfare in recessions than in booms, would survive such a generalization.

Identification of Measurement Error Switching Model

In this Appendix I show how all the parameters of the switching-regression model with coding error (1.1.a)–(1.1.c) are econometrically identified, subject to the normalization that $\sigma_2 = 1$. Recall the definitions

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

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

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

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

Identification of Measurement Error Switching Model

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

$$p_d = D \cdot Prob(D = 1) + (1 - D) \cdot Prob(D = 0) =$$

$$D \cdot \Phi(Z\zeta / \sqrt{1 + \sigma_\eta^2}) + (1 - D) \cdot (1 - \Phi(Z\zeta / \sqrt{1 + \sigma_\eta^2})) \quad (A2.2)$$

$$\pi_i = I \cdot Prob(I = 1) + (1 - I) \cdot Prob(I = 0) = I \cdot \Phi(Z\zeta) + (1 - I) \cdot (1 - \Phi \quad (A2.3)$$

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

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

A Recursion Algorithm

Markovian Switching-Regression Model with Coding Error

The aim is to facilitate evaluation of the likelihood function of Section 4, which is given by:

$$p[\mathbf{y}, \mathbf{D}|X] = \sum_{I_T} \sum_{I_{T-1}} \cdots \sum_{I_2} \sum_{I_1} \sum_{I_0} p[y_T, D_t, I_T | I_{T-1}] \cdots p[y_1, d_1, I_1 | I_0].p[I_0]. \quad (A3.1)$$

A Recursion Algorithm

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.

A Recursion Algorithm

Define the set of available endogenous information at time t by S_t, i.e., $S_t \equiv (y_1, D_1, y_2, D_2, ..., y_t, D_t)$. Further define $Q_t(I_t) \equiv p[S_t, I_t]$. Since we can always write

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

A Recursion Algorithm

it follows that

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

A Recursion Algorithm

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 $Q_{t-1}(I_{t-1}) \equiv p[I_{t-1}, S_{t-1}]$.

A Recursion Algorithm

But calculation of (A3.3) only requires information up to t, as the following matrix equation shows:

A Recursion Algorithm

or,

$$\mathbf{Q}_t = \mathbf{M}_t \cdot \mathbf{Q}_{t-1}.$$

A Recursion Algorithm

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

$$p[\mathbf{y}, \mathbf{D}|X] = \sum_{I_T} Q_T(I_T) = Q_T(0) + Q_T(1). \qquad (A3.1')$$

Figure 1: Conditional Probability $p_{1|0}$

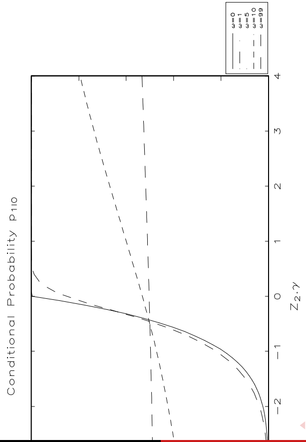

Figure 2: Joint Probability p_{10}

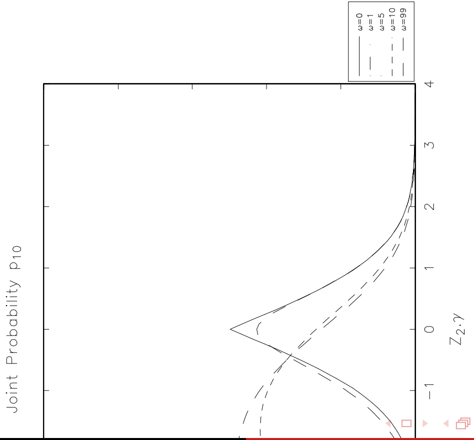

Figure 3: Total Probability p_{10}