An Econometric Analysis of Some Models for Constructed Binary Time Series

Don Harding
Adrian Pagan

An Econometric Analysis of Some Models for Constructed Binary Time Series

Don Harding, Latrobe University, Bundoora, Australia 3056.

Adrian Pagan, University of New South Wales, Sydney Australia, 2052.

and Queensland University of Technology, Brisbane, Australia, 4000.

July 2, 2009

Abstract

Macroeconometric and financial researchers often use binary data constructed in a way that creates serial dependence. We show that this dependence can be allowed for if the binary states are treated as Markov processes. In addition, the methods of construction ensure that certain sequences are never observed in the constructed data. Together these features make it difficult to utilize static and dynamic Probit models. We develop modelling methods that respects the Markov process nature of constructed binary data and explicitly deals with censoring constraints. An application is provided that investigates the relation between the business cycle and the yield spread.

1

Key Words: Business cycle; binary variable, Markov process, Probit model, yield curve

JEL Code C22, C53, E32, E37

1 Introduction

Macroeconometric and financial econometric research often feature discrete random variables that have been *constructed* from some underlying continuous random variable y_t. Often these discrete variables have a binary form, with the two values representing whether an event has occurred or not. An example might be the sign of the change in an interest rate. Such a constructed variable has an *ordinal* nature. There are other cases where the discrete random variable is augmented in such a way as to make it *cardinal* e.g. by identifying the actual values of the signed changes in y_t as the outcomes of the discrete random variable. An influential example would be Eichengreen et al (1985) who implemented such a strategy for the modelling of Bank Rate — the rate of interest charged by the Bank of England to discount houses and other dealers in Treasury bills — as this rate only varied by a small number of discrete movements. If the constructed variables have a cardinal nature then the methods to analyze them are clearly different to the ordinal case, as extra information is available. This paper is concerned with the analysis of ordinal binary variables. The nature of such constructed random variables has not been studied much, notable exceptions being Kedem (1980), Watson (1994), Startz(2008) and Harding and Pagan (2006).

The binary variable we will work with can be thought of as representing the state of some characteristic of the economic and financial system, such as activity or equity market performance. We will designate it as S_t. As

an example data on economic activity can be used to construct a binary variable S_t, taking the value of unity if activity is in an expansion phase and zero when activity is in the contraction phase. Although there are many other examples, including bull and bear markets for stock prices, we will focus mainly upon the case of economic activity i.e. business cycles.

The data S_t is mostly constructed by some individuals and agencies that are external to the researcher. A question that then arises is whether the methods of construction have an impact upon the data generating process (DGP) of the S_t, and, if so, do these features create any special problems for econometric analysis? To make this question more concrete consider the business cycle data available from the NBER. Using the quarterly S_t they present on their web page for 1959/1 to 1995/2 (the same period as in the application we look at in section 5), an OLS regression is run in (1) of S_t on a constant, S_{t-1}, S_{t-2}, and $S_{t-1}S_{t-2}$ (Newey-West HAC t-ratios in brackets for window width four)

$$S_t = \underset{(3.8)}{0.4} + \underset{(5.6)}{0.6S_{t-1}} - \underset{(-3.8)}{0.4S_{t-2}} + \underset{(3.1)}{0.35S_{t-1}S_{t-2}} + u_t \qquad (1)$$

Equation (1) has three striking features. First, the estimated constant term and the coefficient of S_{t-1} sum exactly to one. Second, the estimated constant and the coefficient on S_{t-2} sum exactly to zero. In both cases these results hold exactly independently of the number of decimal places used to

4

represent the coefficients. Third, the process for S_t is at least a second order process. The question is where these features come from. In this paper we show that the method of construction is an important determinant of them and that models should be chosen which recognize that these features will be present.

In section 2 of the paper we explore the interaction between the method of construction of the S_t and the nature of the y_t they are drawn from. We do so by looking at some simple rules for constructing the S_t, which capture the main ways that such variables are constructed. We then interact these rules with a DGP for the y_t chosen so that it mimics the data on the underlying variables from which the S_t derive. The simple examples of this section are meant to aid an understanding of the origin of the features above, and to guide researchers when selecting appropriate models for the S_t.

Often we wish to relate these binary random variables to some regressors x_t. In such circumstances the literature on conditional models for categorical data is mostly employed. Probit and Logit models are well known examples of this class. These relate S_t to a single index $x_t'\theta$, through $\Pr(S_t = 1|x_t) = F(x_t'\theta) = F(z_t)$, where $F(\cdot)$ is a cumulative distribution function with the properties that $F(z)$ is monotonic increasing in z, $\lim_{z \to -\infty} F(z) = 0$ and $\lim_{z \to \infty} F(z) = 1$. The Probit model is the special case where $F(\cdot) = \Phi(\cdot)$, the CDF of the standard normal. Mostly these are static models, as in Estrella and Mishkin (1998), but some dynamic versions have been proposed to handle a time series of categorical data e.g. Kauppi and Saikkonen

5

(2008) and de Jong and Woutersen (forthcoming). In these models, termed a single index dynamic categorical (SIDC) model here, one has $\Pr(S_t = 1 | x_t, S_{t-1}, .., S_{t-p}) = F(x'_t \theta, S_{t-1}, .., S_{t-p})$. It is natural then to seek to utilize those models to describe the DGP of the S_t. Using a popular rule for determining business cycle dates, we show that the DGP of the binary states cannot be represented by the SIDC that has mostly been used. This raises doubts about whether the NBER business cycle states can be represented in this way.

The issue just raised leads in section 3 to a presentation of multiple index generalizations of the SIDC model which can match the features of NBER binary data. These are termed GDC models. In section 4 we adapt a non-parametric estimation method to estimate the GDC model. Section 5 then applies this method to the same sample of NBER data as used by Estrella and Mishkin (1998) when fitting a single index static Probit model. We find that the econometric issues originating from the method of construction of the S_t that have been identified above are empirically significant.

2 The Impact of Method of Construction on the DGP of the Binary Variables

Even though some information might be lost in the process, binary variables are constructed from a primary set of data for at least two purposes. One is to focus attention on particular characteristics. Thus squaring the data

loses information on the sign but emphasizes volatility. In the same way binary random variables locating expansions and recessions focus attention on the frequency and length of these extreme events. The second is to reduce the dimension of the data generating process so as to more easily discern important patterns in the data or to isolate characteristics that a model seeking to interpret the data would need to incorporate. Thus decomposing data such as y_t into its permanent and transitory components is a key step in economic model design. In the same way an interaction of the binary data with the y_t can point to important characteristics such as the rapidity of recovery from an expansion that economic models need to account for.

Often the user of the S_t is not the producer. Consequently, the researcher often just has a set of binary data S_t available and (sometimes) knowledge of the y_t they have been constructed from. To understand the nature of the S_t we therefore need to have some idea of the transformations that link these two series. Although we may not know precisely how this is done, in most instances enough information is provided along with the data on the S_t to enable a good approximation to it. It is worth thinking of the conversion process from y_t to S_t as involving three stages, and to see how the nature (DGP) of S_t changes at each stage. We do this in the subsequent sub-sections.

2.1 Stage 1: Effects of State Change Rules

In the first stage we seek to determine what state the system is in at various points in the sample path. In the business cycle context, where we are

7

seeking states of expansion and contraction, it is often the case that these are identified by locating the *turning points* in the series y_t. Often these first stage turning points are produced by a set of *rules* formalized in algorithms such as that due to Bry and Boschan (BB)(1971) and a simplified quarterly version of it (BBQ) described in Harding and Pagan (2002). In other cases the rules are found by using the output from fitting statistical models such as latent Markov Processes to the y_t series — Hamilton (1989). In all instances these rules transform y_t into S_t.

Because turning point rules are widely used in the analysis of business cycles (and are the basis of the NBER data that we utilize later for empirical work) we focus on them in what follows. Turning points are found by locating the local maxima and minima in the series y_t. A variety of rules appear in the literature to produce the turning points. It is useful to study three of these in order to understand how each rule influences the nature of the univariate DGP for S_t and to understand the inter-relations between S_t and any regressors x_t that are thought to influence the state. The impact of any given rule will also depend upon the DGP of y_t. Consequently, we will study how the mapping between y_t and S_t changes as we modify either the rules or the DGP of y_t.

2.1.1 Calculus rule

The simplest method of locating turning points is what might be termed the *calculus rule*. This says that a *peak* in a series on activity, y_t, occurs

at time t if $\Delta y_t > 0$ and $\Delta y_{t+1} < 0$. The reason for the name is the result in calculus that identifies a maximum with a change in sign of the first derivative from being positive to negative. A *trough* (or local minimum) can be found using the outcomes $\Delta y_t < 0$ and $\Delta y_{t+1} > 0$. The states S_t are simply defined in this case as $S_t = 1(\Delta y_t > 0)$, so that S_t depends only on contemporaneous information. Note that we could formulate this rule as $S_t = 1(\Delta y_t > 0|S_{t-1} = \{0,1\})$ in which case it describes how the state changes, and it might be called a termination rule. This rule has been popular for defining a business cycle when y_t is yearly data, see Cashin and McDermott (2002) and Neftci (1984).

Univariate DGP of S_t Suppose that Δy_t is a Gaussian covariance stationary process and the calculus rule is employed. In this instance, Kedem(1980, p34) sets out the relation between the autocorrelations of the Δy_t and $S(t)$ processes. Letting $\rho_{\Delta y}(k) = corr\left(\Delta y_t, \Delta y_{t-k}\right)$, and $\rho_S(k) = corr\left(S_t, S_{t-k}\right)$, he determines that

$$\rho_S(k) = \frac{2}{\pi} \arcsin\left(\rho_{\Delta y}(k)\right). \tag{2}$$

Thus $corr\left(S_t, S_{t-k}\right) = 0$ only if $corr\left(\Delta y_t, \Delta y_{t-k}\right) = 0$. Notice that the order of the S_t process changes with the degree of serial correlation in the Δy_t series. Since turning points are invariant to monotonic transforms of the data we can think of y_t as being the log of a variable such as activity. Hence the degree of serial correlation in the growth rates of activity will influence the cycle turning points found with the calculus rule.

Relation of S_t and x_t If the underlying process for Δy_t is

$$\Delta y_t = x_t'\theta + \varepsilon_t \qquad (3)$$

where x_t is assumed to be strictly exogenous (and so can be conditioned upon) and ε_t is $n.i.d.(0,1)$. Then $S_t = 1(x_t'\theta + \varepsilon_t > 0)$ and a static Probit model would clearly capture the relation between S_t and the single index $z_t = x_t'\theta$ since $\Pr(S_t = 1|z_t, S_{t-1}) = \Phi(z_t)$.

2.1.2 Two quarters rule

The rule that two quarters of negative growth terminates a recession is often cited in the media. Extended so that the start of an expansion is identified with two quarters of positive growth produces the "two quarters rule":

$$
\begin{aligned}
S_t &= 1 \text{ if } (\Delta y_{t+1} > 0, \Delta y_{t+2} > 0 | S_{t-1} = 0). \\
S_t &= 0 \text{ if } 1(\Delta y_{t+1} < 0, \Delta y_{t+2} < 0 | S_{t-1} = 1) \qquad (4) \\
S_t &= S_{t-1} \text{ otherwise.}
\end{aligned}
$$

Lunde and Timmermann (2004) used a variant of this non-parametric rule for finding bull and bear periods in stock prices while hot and cold markets for IPO's were identified by Ibbotson and Jaffee (1975), with a hot market being signalled by whether excess returns and their changes for two periods

exceed the median values. Eichengreen et al. (1995) and Classens et al (2008) employ rules of this type to establish the location of crises in time.

Univariate DGP of S_t To illustrate the features of this rule consider the case where y_t is a Gaussian random walk with drift

$$\Delta y_t = \mu + \sigma e_t, \tag{5}$$

where e_t is $n.i.d(0,1)$ and $\Pr(\Delta y_t < 0) = \Phi\left(\frac{-\mu}{\sigma}\right) \equiv \psi$. Then we show in the appendix that the first order representation of S_t has the following parameterization

$$S_t = \frac{(1-\psi)^2}{2-\psi} + [1 - \frac{(1-\psi)^2}{2-\psi} - \frac{\psi^2}{(1+\psi)}]S_{t-1} + \eta_t, \tag{6}$$

This example shows that even where y_t is a Gaussian random walk the use of a two quarters rule rather than the calculus rule induces serial correlation into the S_t process so that it is at least a first order process.

Relation of S_t **and** x_t It is also instructive to examine the case where Δy_t has the DGP in (3). The appendix derives $\Pr(S_t = 1|S_{t-1} = 1, x_t, x_{t-1}, ...)$, and it is found to depend non-linearly upon the complete history $\{x_{t-j}\}_{j=0}^{\infty}$, with the non-linear mapping failing to be that provided by the CDF of an $N(0,1)$ variable. Hence a static Probit model would be inappropriate. A dynamic one, in which lags of S_t are added to the single index, would certainly

imply dependence of the probability on past values of x_t, but the non-linear mapping between S_t and $\{x_{t-j}\}_{j=0}^{\infty}$ would be incorrect. For this reason we need to allow for a general functional relation connecting S_t and x_t, and we therefore set out methods for doing this in the next section.

2.1.3 Bry-Boschan and BBQ rules

Neither the calculus rule nor the "two quarters" rule accurately describes the rule used by the NBER to locate local peaks and troughs in y_t. To match the features of that data requires a rule that formalizes the visual intuition that a local peak in y_t occurs at time t if $y_t > y_s$ for s in a window $t - k < s < t + k$ — a trough is defined in a similar way. By making k large enough we also capture the idea that the level of activity has declined (or increased) in a sustained way. This rule with $k = 5$ months is the basis of the NBER business cycle dating procedures summarized in the Bry and Boschan (1971) dating algorithm. The comparable BBQ rule sets $k = 2$ for quarterly data. These turning point rules have been used in other contexts than the business cycle e.g. the dating of bull and bear markets in equity prices by Pagan and Sussonov (2003), Bordo and Wheelock (2006) and Claessens et al (2008).

It seems very difficult to analytically determine what the impact of these rules would be upon the order of serial correlation in the S_t. Simulations however show that the results mimic those found with the two-quarter-rule, so that this is quite a good guide to what one might expect if NBER dating

methods are employed.

2.1.4 Markov processes

As seen above the order of the univariate process for S_t and the relation between between S_t and x_t varies with the dating rule and the nature of x_t. This suggests that we need to keep both the order and any functional relation as flexible as possible and raises the issue of what type of representation we might want for S_t. When seeking general representations of binary time series it is natural to apply the folk theorem (see Meyn 2007, p538) that "every process is (almost) Markov". In our context this would mean that S_t will follow processes like (1), which we will term the Markov process of order two (MP(2)). Higher order MP's would involve higher order lags and cross products between the lagged values. Because these MP processes are effectively non-linear autoregressions they can approximate processes such as Startz's (2008) (Non-Markov) Binary ARMA (BARMA) model to an arbitrary degree of accuracy provided they are of sufficiently high order. Just as VAR's are mostly preferred to VARMA processes in empirical work due to their ease of implementation, we feel that Markov processes should be the work horse when modelling binary time series. They also provide a guide to how one would extend the model linking S_t and x_t, a topic we take up in the next section.

2.2 Stage 2: Effects of State Duration Rules

The second stage in constructing S_t from y_t involves selecting turning points that satisfy certain requirements related to minimum completed phase lengths. This process is referred to here as "censoring" and it is evident in many data series on S_t. It is designed to ensure that once a state is entered it persists for some time. So recessions and expansions or crises should continue for a certain minimum period of time. In this context the standard requirement of the NBER when dating business cycles is that completed phases have a duration of at least two quarters. This requirement is evident in the NBER data — from its beginning in 1859 onwards there is no completed phase with duration of less than two quarters.

2.2.1 Univariate DGP of S_t

Using the NBER censoring restrictions just noted, any DGP for S_t must have the properties that

$$\Pr\left(S_t = 1 | S_{t-1} = 1, S_{t-2} = 0\right) = 1 \tag{7}$$

and

$$\Pr\left(S_t = 1 | S_{t-1} = 0, S_{t-2} = 1\right) = 0. \tag{8}$$

Now we have suggested that the S_t be treated as an MP. Suppose it has

the form of the MP(2) in (1) viz.

$$S_t = \alpha_0 + \alpha_1 S_{t-1} + \alpha_2 S_{t-2} + \alpha_2 S_{t-1} S_{t-2} + u_t, \tag{9}$$

where $E\left(u_t | S_{t-1}, S_{t-2}\right) = 0$. Now, for binary data,

$$\Pr\left(S_t = 1 | S_{t-1} = s_1, S_{t-2} = s_2\right) = E\left(S_t | S_{t-1} = s_1, S_{t-2} = s_2\right) \tag{10}$$

and the properties (7) and (8) imply the parameter restrictions that

$$\alpha_0 + \alpha_2 \;=\; 0 \tag{11}$$

$$\alpha_0 + \alpha_1 \;=\; 1. \tag{12}$$

This establishes that the empirical features identified in the introduction to the paper are indeed directly caused by the censoring process used by the NBER. Notice that this is independent of the turning point rule used, so that the order of the serial correlation in the S_t process may also stem simply from a censoring procedure.

2.2.2   Relation of S_t and x_t

We now turn to the issue of the implications of the properties (7) and (8) for SIDC models. As noted in the introduction it has often been the case that SIDC models have been constructed by a mapping between S_t and a single index made up of $x_t'\theta$ and lags of S_t. In the absence of x_t this would

be expected to have the form in (9) and this suggests that the appropriate generalization of the SIDC model (of second order) would be

$$
\begin{aligned}
z_t \; = \; & x_t'\theta + \beta_1 \left(1 - S_{t-1}\right)\left(1 - S_{t-2}\right) + \beta_2 \left(1 - S_{t-1}\right) S_{t-2} \qquad (13) \\
& + \beta_3 S_{t-1} \left(1 - S_{t-2}\right) + \beta_4 S_{t-1} S_{t-2},
\end{aligned}
$$

where $\Pr(S_t = 1 | z_t) = F(z_t)$.

The properties (7) and (8) that are attributable to the censoring of the states in order to achieve minimum phase duration also impose restrictions on the parameters of the models linking S_t and x_t. Specifically,

$$
\Pr\left(S_t = 1 | S_{t-1} = 1, S_{t-2} = 0, x_t\right) = 1 = F\left(x_t'\theta + \beta_3\right) \qquad (14)
$$

and

$$
\Pr\left(S_t = 1 | S_{t-1} = 0, S_{t-2} = 1, x_t\right) = 0 = F\left(x_t'\theta + \beta_2\right). \qquad (15)
$$

Since $F\left(\cdot\right)$ is a CDF the true parameter values required to satisfy (14) and (15) are $\beta_3 = \infty$ and $\beta_2 = -\infty$, which violates the standard regularity conditions for an MLE estimator viz. that the parameter space be a compact set and the maximum be in the interior of this. Note that this problem arises because of the inclusion of terms involving S_{t-2} in the functional form linking S_t and x_t. It would not have arisen had we used $z_t = x_t'\theta + \beta_1 S_{t-1}$ only. But this would amount to disregarding the fact that S_t must be a second order process whenever the available data has been censored. Clearly dynamic

models for the binary time series must be developed that adapt to the order of dynamics of the binary variables, and we return to that in the next section.

2.3   Stage 3: Judgement

Although there are exceptions, in most instances the S_t researchers are presented with involve modifying the S_t that would one would get from the two stages above. This modification stems from the application of expert judgement. It should be emphasized that there is no doubt that the two stages above are inputs into the final decision. Accordingly, the lessons learned from the analysis presented above are important for working with the final S_t. In particular, the nature of the process for S_t established in stages one and two is likely to carry over to the final states. This is evident from (1), where the S_t used in the regression are the final states selected by the NBER Dating Committee. It has also been found that there is a close correspondence between the published NBER S_t and those coming from an application of the BB and BBQ algorithms. In many ways the situation is like a Taylor rule for describing interest rate decisions. The FOMC do not use a linear Taylor rule but it is often a good description of their behavior. But one should be wary of assuming that it is a precise description. It may be that the information in the Taylor rule maps into the decision in a non-linear way or with a different lag structure. Thus one needs to be flexible in how one models these decisions. Analogously, we cannot utilize the results derived in the previous section to give precise models that could be fitted to the S_t, but

rather the models suggested by our analysis of rules and censoring provide essential guidance on what might be sensible models to entertain.

3 Generalized Dynamic Models for Binary Variables

Since the S_t are binary variables $E(S_t|S_{t-1}, S_{t-2}, x_t)$ is $F(z_t)$ in (13). It is useful to re-parameterize it as

$$
\begin{aligned}
F(z_t) = \ & F\left(x_t'\theta + \beta_1\right)\left(1 - S_{t-1}\right)\left(1 - S_{t-2}\right) + F\left(x_t'\theta + \beta_2\right)\left(1 - S_{t-1}\right)S_{t-2} \\
& + F\left(x_t'\theta + \beta_3\right)S_{t-1}\left(1 - S_{t-2}\right) + F\left(x_t'\theta + \beta_4\right)S_{t-1}S_{t-2}.
\end{aligned}
\tag{16}
$$

Now this form came from using a model for y_t that had errors with a CDF of the form $F(z_t)$. But this might be regarded as a strong restriction since the investigations reported in the previous section suggested that it is unlikely that a mapping of this sort between S_t and x_t will obtain. It is desirable to let the data determine what the functional relation is, provided that our generalized model nests that in (16). Thus we express $\Pr(S_t = 1|S_{t-1} = s_1, S_{t-2} = s_2, x_t)$ as $G\left(s_1, s_2, x_t\right)$, where

$$
\begin{aligned}
G\left(S_{t-1}, S_{t-2}, x_t\right) = \ & \beta_1\left(x_t\right)\left(1 - S_{t-1}\right)\left(1 - S_{t-2}\right) + \beta_2\left(x_t\right)\left(1 - S_{t-1}\right)S_{t-2} \\
& + \beta_3\left(x_t\right)S_{t-1}\left(1 - S_{t-2}\right) + \beta_4\left(x_t\right)S_{t-1}S_{t-2}.
\end{aligned}
\tag{17}
$$

18

In the representation (17) $\beta_i(x_t)$ are functions with the property that $0 \leq \beta_i(x_t) \leq 1$ but there is now no longer the requirement that the $x's$ be formed into a single index, nor is there a requirement that they take a particular functional form. The censoring requirements embedded in properties (7) and (8), however, do require that

$$\beta_2(x_t) = 0$$

and

$$\beta_3(x_t) = 1.$$

Consequently, with these restrictions in place,

$$G(S_{t-1}, S_{t-2}, x_t) = \beta_1(x_t)(1 - S_{t-1})(1 - S_{t-2}) + S_{t-1}(1 - S_{t-2}) + \beta_4(x_t)S_{t-1}S_{t-2}.$$
(18)

In general (18) is a two index model. It only becomes a single index model in the special case where $\beta_1(x_t) = \pm\beta_4(x_t)$. Thus the single index restriction is a hypothesis that is readily tested. We will refer to the model in (17) as a generalized Dynamic Categorical model (GDC)

4 Non-parametric Estimation of the GDC Model

The model involves estimating $m_c = E(S_t|S_{t-1}, S_{t-2}, x_t)$ from (17) or its censored version (18). Because we will want to compare these models to

19

others that have been applied, such as static Probit, which only evaluate the expectation of S_t conditioned upon x_t, we will need to find an expression for the $E(S_t|x_t)$ that is implicit in them. To do that requires S_{t-1} and S_{t-2} to be integrated out of the conditional expectation m_c. Doing so yields

$$
\begin{aligned}
E(S_t|x_t) & = \sum_{j=0}^{1}\sum_{k=0}^{1} E(S_t|S_{t-1}=j, S_{t-2}=k, x_t) & (19) \\
\times \Pr(S_{t-1} & = j, S_{t-2}=k|x_t) \\
& = \sum_{j=0}^{1}\sum_{k=0}^{1} m_c(j,k,x_t) \Pr(S_{t-1}=j, S_{t-2}=k|x_t) & (20)
\end{aligned}
$$

Consequently, to find $E(S_t|x_t=x)$ we will need to estimate the conditional expectations

$$
\begin{aligned}
m_c(1,1,x) & = E(S_t|S_{t-1}=1, S_{t-2}=1, x) = \beta_4(x) & (21) \\
m_c(1,0,x) & = E(S_t|S_{t-1}=1, S_{t-2}=0, x_t) = \beta_3(x) & (22) \\
m_c(0,1,x) & = E(S_t|S_{t-1}=0, S_{t-2}=1, x_t) = \beta_2(x) & (23) \\
m_c(0,0,x) & = E(S_t|S_{t-1}=0, S_{t-2}=0, x_t) = \beta_1(x). & (24)
\end{aligned}
$$

Once the conditional expectations are found estimates of $\beta_j(x)$ can be extracted. Accordingly, we focus upon methods for estimating the conditional expectations $m_c(j,k,x) = E(S_t|S_{t-1}=j, S_{t-2}=k, x_t=x)$ by nonparametric methods, specifically a kernel estimator.

Now in our application of the next section there will be two categorical variables (S_{t-1}, S_{t-2}) and one variable which is likely to be continuous (x_t).

Estimation of a conditional expectation involving such variables by kernel methods has been extensively discussed in Racine and Li (2004). A multivariate kernel will need to be used and we will follow the standard practice of making this the product of univariate ones, so that three univariate kernels are needed for each of the conditioning variables. Racine and Li (2004) propose that the kernel used for a categorical variable z_t take the value $K(z_t, z, \lambda) = 1$ when $z_t = z$, and the value λ otherwise. They suggest that, when the categorical variable takes a number of values which are not well differentiated, λ be estimated. Otherwise a value of $\lambda = 0$ is satisfactory. A value of $\lambda = 0$ would mean that the kernel is the indicator function $1(z_t = z)$. Given that our categorical values are binary it seems reasonable to use the indicator function as their kernels, and to adopt a different form, $K(\frac{x_t - x}{h})$, for the continuous random variable. Based on these arguments the estimator of the conditional expectation will be

$$
\begin{aligned}
\hat{m}_c(s_1, s_2, x) &= \frac{\sum_{t=1}^{T}\{1(S_{t-1} = s_1) \times 1(S_{t-2} = s_2) \times K(\frac{x_t - x}{h})\}S_t}{\sum_{t=1}^{T}\{1(S_{t-1} = s_1) \times 1(S_{t-2} = s_2) \times K(\frac{x_t - x}{h})\}} \\
&= \frac{\sum_{t \in I_{jk}} S_t K(\frac{x_t - x}{h})}{\sum_{t \in I_{jk}} K(\frac{x_t - x}{h})},
\end{aligned}
$$

where I_{jk} are those observations for which $S_{t-1} = j, S_{t-2} = k$. Racine and Li (2004) show that, under the assumption that S_t and x_t are independently

21

distributed, and h varies with T in a standard way, the estimator $\hat{m}_c(j, k, x)$ is a consistent estimator of $m_c(j, k, x)$, and the following central limit theorem holds

$$\sqrt{Th}(\hat{m}_c(j, k, x) - m_c(j, k, x)) \rightarrow \mathcal{N}(0, \frac{\int K(\psi)^2 \, d\psi}{f(x)} \sigma_\eta^2(j, k, x)). \qquad (25)$$

In (25)

$$
\begin{aligned}
\sigma^2(j, k, x) &= E\left[\eta_t^2 | S_{t-1} = j, S_{t-2} = k, x\right] \\
&= m(j, k, x)(1 - m(j, k, x)),
\end{aligned}
$$

with the latter coming from the binary nature of S_t.

Now our situation differs from that described above since S_t and x_t are unlikely to be independently distributed. However, Li and Racine (2007, Theorem 18.4) extend this result to the case that η_t is a martingale difference process and x_t, S_t are stationary $\beta-$mixing processes. This is in line with earlier results by Bierens (1983) and Robinson (1983). We will therefore assume that such conditions are satisfied for S_t and x_t.

Turning to $\Pr(S_t = 1 | x_t = x) = m(x) = E(S_t | x_t = x)$ it follows from (19) that

$$(Th)^{\frac{1}{2}}(\hat{m}(x) - m(x)) \rightarrow \mathcal{N}\left(0, \frac{\int K(\psi)^2 \, d\psi}{f(x)} \sum_{j=0}^{1} \sum_{k=0}^{1} \sigma_\eta^2(j, k) \Pr(t \in I_{jk} | x)\right).$$

$$(26)$$

In the application of the next section we use (25) and (26) to establish asymptotic confidence intervals for the non-parametric estimators of the requisite expectations.

It is useful to note that if one is interested simply in $E(S_t|x_t = x)$ then the kernel density estimator (27) is exactly equivalent to (20) — a formal proof is available from the authors on request. Of course the standard errors from (27) are incorrect and this is one reason to favour (20). Also, as is demonstrated in the application, the intermediate results used to calculate (20) are of considerable interest in their own right.

$$E(S_t|x_t) = \frac{\sum_{t=1}^{T} S_t K\left(\frac{x_t - x}{h}\right)}{\sum_{t=1}^{T} K\left(\frac{x_t - x}{h}\right)} \qquad (27)$$

# 5 An Application to the Probability of Recessions Given the Yield Spread

We apply the methods developed above to assess the extent to which the yield spread (sp_t) affects the probability of a recession occurring. Estrella and Mishkin (1998) assessed this question by applying a static Probit model to the NBER states i.e. a Probit model was assumed to give a functional relation between S_t and the spread. This amounts to ignoring the dependence in and censoring of the binary variable S_t. In this application we use the methods

described earlier to take account of the fact that the NBER states S_t are neither independent, identically distributed nor uncensored. The conditional mean is calculated using (19) with a kernel that is a product of Gaussian densities.

Estrella and Mishkin (1998) find that the best fit occurs with the yield spread being lagged two quarters, and we continue with that assumption here, so that $x_t = sp_{t-2}$ in this application. Figure 1 plots the probability of a recession given the spread i.e. $E(1 - S_t|sp_{t-2})$, against sp_{t-2} found in two ways. One is by estimating a static Probit model and the other is the implied value coming from first fitting the GDC model in (17) to the data and then using (20) to compute the requisite expectation. Before doing so a test was performed on whether the Markov process for S_t should be third rather than second order and the latter was favoured. Also shown on the figure are the 95% confidence bands obtained using the asymptotic results for the estimator of $1 - E(S_t|x_t)$ given in (26). It is clear that there is a difference between the probability of recession obtained from the static Probit and GDC model at a number of values for the spread. Most notably this occurs for spreads in the range -0.55% to 0.5%, although there is close to being a significant difference for a spread around -1% ; at that point the static Probit model yields a predicted probability of recession that is much lower than the GDC model.

Having established that making an allowance for the nature of S_t is both theoretically and empirically important it is of interest to evaluate the extent

24

Figure 1: Probability of recession from MP(2) and Probit models conditional on the yield spread lagged two quarters
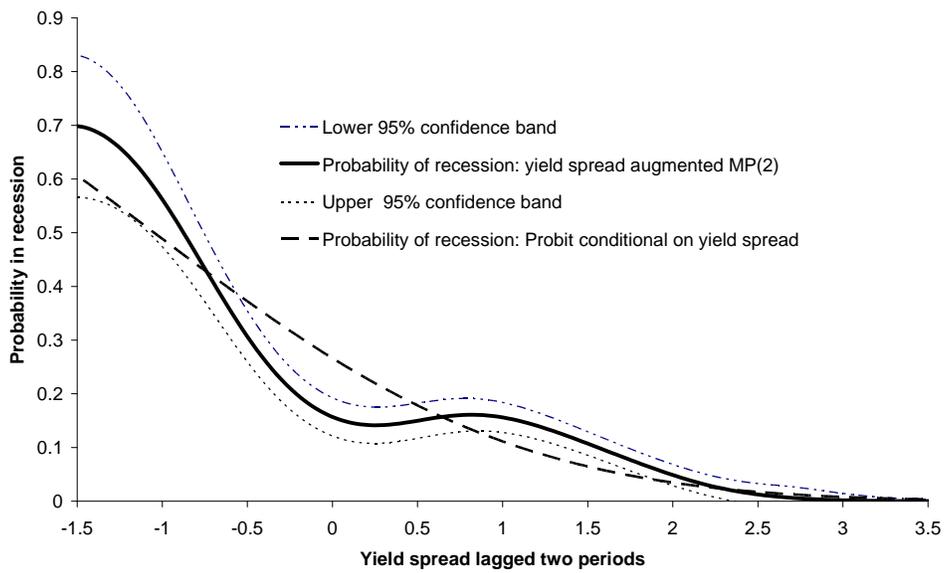
to which the yield spread is useful when looking at the probability of moving from an established phase to the opposite one. To assess this we focus on either the probability that an expansion which has lasted for two or more periods will be terminated or the probability of continuing in a contraction that has lasted for two or more periods. The former is the quantity $E(S_t = 0 | x_t, S_{t-1} = 1, S_{t-2} = 1)$, while the latter is $E(S_t = 1 | x_t, S_{t-1} = 0, S_{t-2} = 0)$.

The probability of leaving an expansion that has lasted for two or more quarters is shown in Figure 2. There is a substantial difference between the estimates obtained from the non-parametric estimates of the GDC model and those from a *dynamic* Probit model that uses sp_{t-2} and S_{t-1} as covariates (the variant used in some of the cycle literature). The dynamic Probit model over-predicts the probability of leaving an expansion for yield spreads in the range -1.1% to 0.3%, and under-predicts the probability of leaving an expansion for yield spreads below -1.1%. These differences are statistically significant at the 5% level for spreads in the interval -0.75% to 0%. The main finding, that the probability of terminating an expansion is low for spreads above -0.75, should be of interest to policy makers.

The probability of continuing in a recession that has lasted for two quarters is plotted in Figure 3. Again the probabilities are from the GDC model and the dynamic Probit model. There is a substantial difference between the predicted probabilities from the two models, and this difference is both economically and statistically significant. The most important difference between the probabilities from the two methods is that the GDC model suggests

Figure 2: Probability of terminating an expansion that has lasted for two quarters
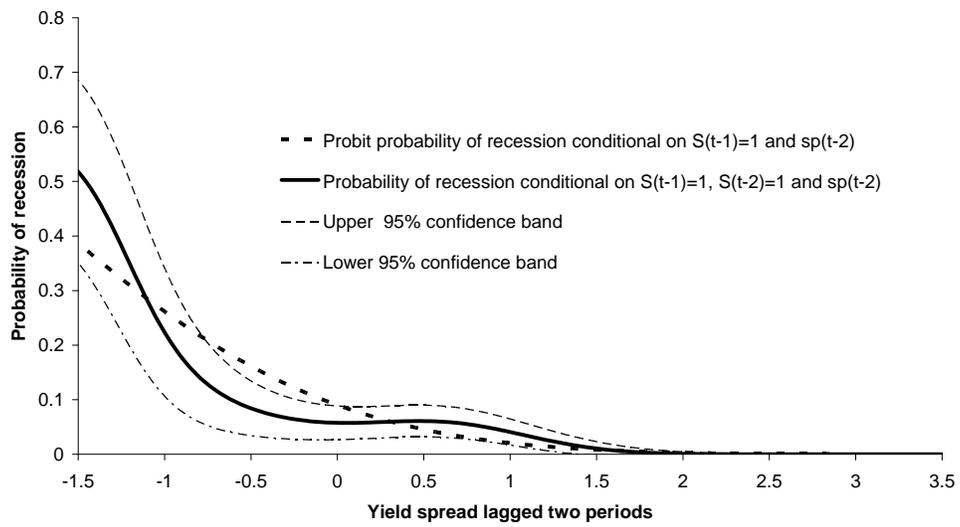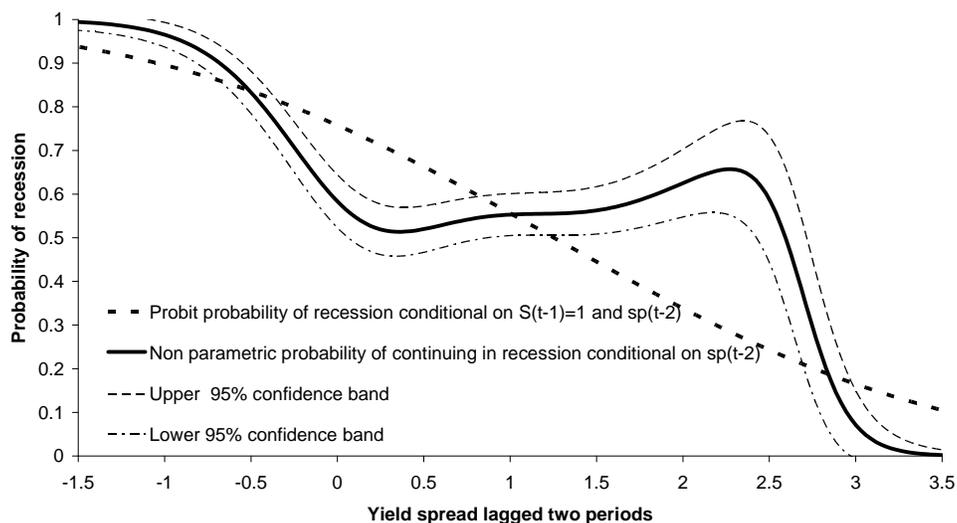
Figure 3: Probability of continuing in a recession

that there is no decrease in the probability of staying in a recession, with a rise in the yield spread from zero to 2.5 per cent. In contrast the dynamic Probit model suggests that the probability of remaining in recession declines monotonically as the yield spread increases.

Of course, one may question the accuracy of the asymptotic confidence intervals for this experiment, as there are only 10 per cent of cases where the economy is in contraction for two or more periods. But, even allowing for this caveat, the results presented above are likely to be of considerable practical interest.

6 Conclusion

We have argued that constructed states S_t require careful treatment if they are to be used in econometric work, since they are very different in their nature to the binary states often modelled in micro-econometrics. When engaging in a broad range of estimation and inference methods one has to allow for the fact that they are essentially Markov processes. But, to date, the nature of the S_t has mostly been ignored, with the potential for misleading estimates and inferences. We have suggested some methods to deal with this fact. In the application these methods produce results that differ from those obtained by a standard Probit procedure that does not allow for the Markov process nature of the binary states and which forces a particular functional form upon the data. We have shown that these differences are economically and statistically significant.

Appendix A: Obtaining Transition Probabilities Under The "Two Quarters Rule"

The task of obtaining transition probabilities becomes much more complex with the "two quarters rule" as the conditioning event $S_{t-1} = 1$ will place some restrictions upon the past sample paths for $\{\Delta y_t\}$ that can be associated with the transition from an expansion to contraction. In this appendix we first set out a procedure for enumerating sample paths that are consis-

tent with $S_{t-1} = 1$. We then apply that procedure to obtain the univariate transition probabilities when y_t follows a random walk with drift. We first investigate the case where the drift is a constant and then study the case where the drift is a function of some exogenous random variable x_t. With the "two quarters" rule the key feature of the sample path is whether Δy_t is positive or negative. We use "$+_t$" to denote the former and "$-_t$" to denote the latter.

Enumerating sample paths

Using the "$+_t$" "$-_t$" notation the sequence for $sign\{\Delta y_t\}$

$$\{-_{t+1}, -_t, -_{t-1}, +_{t-2}, ...\} \tag{28}$$

would be incompatible with $S_{t-1} = 1$ since the negative growth at $t-1$ would match with the negative growth at t and so $t-2$ would be the last period of an expansion thereby making $t-1$ the first period of a contraction.

From this example it is clear that the sample paths $\{\Delta y_{t-1}, \Delta y_{t-2}, ...\}$ that are compatible with both $S_{t-1} = 1$ and $\{\Delta y_{t+1} < 0, \Delta y_t < 0\}$ must have positive growth at $t-1$. Moreover, in such paths we must encounter a $\{+, +\}$ before we encounter a $\{-, -\}$. If this did not happen so that, for example, we had the path $\{+_{t-1}, -_{t-2}, +_{t-3}, -_{t-4}, -_{t-5}, ...\}$, then the economy would have been in contraction at $t-5$ and would still be in contraction, according to the two quarters rule, when we reach $t-1$

30

Now let us consider an enumeration of the paths that are consistent with $S_{t-1} = 1$. This is done in the table 1 below where the first column represents time and subsequent columns represent paths along which we are assured that $S_{t-1} = 1$; we have numbered these paths with integers starting with one. The notation used in table 1 is as follows:

- "∗" before a "−" indicates that any pattern for the observations can occur along the path up to and including that point;

- "∗" following a "+" indicates that any pattern for the observations can occur along the path from that point forward.

Thus looking at the second column in table 1 the "+, +" at t and $t-1$ assures us that for path 1 it is the case that $S_{t-1} = 1$.

Turning to path 2, the "−" at t and the "+, +" at $t-1$ and $t-2$ assures us that all paths with this pattern are consistent with $S_{t-1} = 1$. Similar logic can be applied to all the subsequent paths.

To understand the derivation of these paths suppose we start with the four possible outcomes for $(\Delta y_t, \Delta y_{t-1})$, namely $\{+_t, +_{t-1}\}, \{-_t, +_{t-1}\}, \{+_t, -_{t-1}\}$ and $\{-_t, -_{t-1}\}$. The last of these pairs requires that $S_{t-1} = 0$ and the first requires that $S_{t-1} = 1$; thus the first pair becomes the second column of the table. The other two outcomes do not enable us to decide what the state for S_{t-1} is and so we proceed to observation $t-2$ and consider what

31

Table 1: Enumerated paths consistent with being in expansion at time t-1

Time	Path number						
	1	2	3	4	5	6	···
$t+1$	*	*	*	*	*	*	···
t	+	−	+	−	+	−	···
$t-1$	+	+	−	+	−	+	···
$t-2$	*	+	+	−	+	−	···
$t-3$		*	+	+	−	+	···
$t-4$			*	+	+	−	···
$t-5$				*	+	+	···
$t-6$					*	+	···
⋮						*	···

happens to each of them as we add on a $-$ or a $+$. Thus $\{-,+,+\}$ will give $S_{t-1} = 1$ and that becomes the third column. But $\{-,+,-\}$ produces no resolution and one needs to proceed to $t-3$. Augmenting $\{+,-\}$ with a $+$ also fails to resolve the indeterminacy while adding on a $-$ results in $S_{t-1} = 0$. Consequently that path has to be continued on to $t-3$ as well. The process continues in this way and all columns of the matrix will eventually be enumerated by such a strategy.

To formalize the discussion it is helpful to separate the set of paths that are consistent with $S_{t-1} = 1$ into two subsets. Let E_t be the set of paths such that $\{\Delta y_t > 0 \;\; and \;\; S_{t-1} = 1\}$ and F_t be the set of paths such that $\{\Delta y_t < 0 \;\; and \;\; S_{t-1} = 1\}$. If we introduce the notation that

- $[+-]_t^j$ represents the fragment of the path along which there are j repetitions of the pattern $[+-]$ with the leading term in the pattern being located at time t,

- $[++]_t$ represents the fragment of path where the pattern "$++$" occurs with the first "$+$" being at t and the second at $t-1$

- $[-]_t$ represents the case where $\Delta y_t < 0$,

the sets E_t and F_t can be enumerated as

$$E_t = \left\{ [++]_t \, ; [+-]_t \, [++]_{t-2} \, ; [+-]_t^2 \, [++]_{t-4} \, ; \, ...; [+-]_t^j \, [++]_{t-2j} \, ; \, ... \right\} \quad (29)$$

$$F_t = \left\{ \begin{array}{c} [-]_t \, [++]_{t-1} \, ; [-]_t \, [+-]_{t-1} \, [++]_{t-3} \, ; \\ [-]_t \, [+-]_{t-1}^2 \, [++]_{t-5} \, ; \, ...; [-]_t \, [+-]_{t-1}^j \, [++]_{t-2j-1} \, ; \, ... \end{array} \right\}. \quad (30)$$

Interest centres on the joint event $\{S_t = 0, S_{t-1} = 1\}$ that defines a shift from expansion to contraction phase. This will involve the set G_{t+1} that is enumerated as

$$G_{t+1} = \left\{ \begin{array}{c} [--]_{t+1} \, [++]_{t-1} \, ; [--]_{t+1} \, [+-]_{t-1} \, [++]_{t-3} \, ; [--]_{t+1} \, [+-]_{t-1}^2 \, [++]_{t-5} \, ; \, ... \\ ...; [--]_{t+1} \, [+-]_{t-1}^j \, [++]_{t-2j-1} \, ; \, ... \end{array} \right\}$$

$$(31)$$

33

Transition probabilities when y_t is a Gaussian random walk with constant drift

Here $\Delta y_t = \mu + \sigma e_t$ where $e_t \sim N(0,1)$ and $\psi_t \equiv \Pr(\Delta y_t \leq 0) = \Phi\left(\frac{-\mu}{\sigma}\right)$ where $\Phi(\cdot)$ is the CDF of the standard normal.

Thus, using the notation that $\Pr(E_t)$ represents the probability that the path is drawn from the set E_t, and recognizing that the sets E_t and F_t are mutually exclusive, we have,

$$\Pr(E_t) = \sum_{j=0}^{\infty} \Pr\left([+-]_t^j [++]_{t-2j}\right) \tag{32}$$

and

$$\Pr(F_t) = \sum_{j=0}^{\infty} \Pr\left([-]_t [+-]_{t-1}^j [++]_{t-2j-1}\right). \tag{33}$$

By virtue of the definition of E_t and F_t

$$\Pr(S_{t-1} = 1) = \Pr(E_t) + \Pr(F_t). \tag{34}$$

Combining the above results the probability of transiting from expansion to contraction $p_{10} \equiv \Pr(S_t = 0 | S_{t-1} = 1)$ is defined as

$$p_{10} \equiv \frac{\Pr(S_t = 0, S_{t-1} = 1)}{\Pr(S_{t-1} = 1)} = \frac{\Pr(G_{t+1})}{\Pr(E_t) + \Pr(F_t)}. \tag{35}$$

The assumption that y_t is a random walk with constant drift and variance

implies that

$$
\begin{aligned}
\Pr(E) &= \sum_{j=0}^{\infty} (1-\psi)^2 \left[\psi(1-\psi)\right]^j \\
&= \frac{(1-\psi)^2}{1-\psi(1-\psi)} \tag{36}
\end{aligned}
$$

$$
\begin{aligned}
\Pr(F) &= \sum_{j=0}^{\infty} \psi\,(1-\psi)^2 \left[\psi\,(1-\psi)\right]^j \\
&= \frac{\psi(1-\psi)^2}{1-\psi(1-\psi)} \tag{37}
\end{aligned}
$$

So,

$$
\Pr\left(S_{t-1} = 1\right) = \frac{(1+\psi)\,(1-\psi)^2}{1-\psi(1-\psi)} \tag{38}
$$

and

$$
\begin{aligned}
\Pr\left(G\right) &= \sum_{j=0}^{\infty} \psi^2\,(1-\psi)^2 \left[\psi\,(1-\psi)\right]^j \\
&= \frac{\psi^2\,(1-\psi)^2}{1-\psi\,(1-\psi)}. \tag{39}
\end{aligned}
$$

Combining (35), (36), (37) and (39) yields

$$
p_{10} = \frac{\psi^2}{(1+\psi)}. \tag{40}
$$

35

Since S_t is a stationary process $\Pr(S_t = 1, S_{t-1} = 0)$ and $\Pr(S_t = 0, S_{t-1} = 1)$ are constant and, since turning points alternate, $\Pr(S_t = 1, S_{t-1} = 0) = \Pr(S_t = 0, S_{t-1} = 1) = \frac{\psi^2(1-\psi)^2}{1-\psi(1-\psi)}$. Thus

$$p_{01} = \frac{(1 - \psi)^2}{2 - \psi}. \tag{41}$$

The probabilities that the economy stays in the same state, i.e. p_{00} and p_{11} are found from the identities

$$1 = p_{10} + p_{11} \tag{42}$$

and

$$1 = p_{01} + p_{00}, \tag{43}$$

yielding

$$p_{11} = \frac{1 + \psi - \psi^2}{(1 + \psi)} \tag{44}$$

and

$$p_{00} = \frac{1 + \psi - \psi^2}{2 - \psi}. \tag{45}$$

Finally, the transition probabilities can be combined into the following first order equation

$$S_t = p_{01} + [p_{11} + p_{01}] S_{t-1} + \nu_t \tag{46}$$

Transition probabilities when y_t is a random walk with time varying drift

Now in some of the literature we deal with it is assumed that the process for Δy_t depends linearly upon some other variable x_t in the following way:

$$\Delta y_t = a + bx_t + \varepsilon_t, \tag{47}$$

where the x_t are taken to be strictly exogenous (and so can be conditioned upon) and ε_t is $n.i.d.(0,1)$. It will be convenient to let $\Im_t = \{x_{t-i}\}_{i=0}^{\infty}$ represent the history of this exogenous variable.

If $\psi_t = \Phi\left(-a - bx_t\right),$ applying the enumeration method results in

$$
\begin{aligned}
\Pr\left(E_{t+1}|\Im_{t+1}\right) &= \sum_{j=0}^{\infty} \Pr\left([+-]_{t+1}^{j}[++]_{t+1-2j}\right) \\
&= \left(1 - \psi_{t+1}\right)\left(1 - \psi_t\right) \tag{48} \\
&\quad + \sum_{j=1}^{\infty}\left\{\left[\prod_{i=0}^{j-1}\left(1 - \psi_{t+1-i}\right)\psi_{t-i}\right]\left(1 - \psi_{t-2j+1}\right)\left(1 - \psi_{t-2j}\right)\right\}
\end{aligned}
$$

and

$$\begin{aligned}
\Pr\left(F_{t+1}|\Im_{t+1}\right) &= \sum_{j=0}^{\infty} \Pr\left([-]_{t+1}[+-]_t^j[++]_{t-2j}\right) \\
&= \psi_{t+1}\left(1-\psi_t\right)\left(1-\psi_{t-1}\right) \\
&\quad +\psi_{t+1}\sum_{j=1}^{\infty}\left\{\left[\prod_{i=0}^{j-1}\left(1-\psi_{t-i}\right)\psi_{t-i-1}\right]\left(1-\psi_{t-2j}\right)\left(1-\psi_{t-2j-1}\right)\right\}.
\end{aligned} \tag{49}$$

The sets E_{t+1} and F_{t+1} are mutually exclusive and encompass all of the paths along which $S_t = 1$ under the two quarters rule. Thus,

$$Pr\left(S_t = 1|\Im_{t+1}\right) = \Pr\left(E_{t+1}|\Im_{t+1}\right) + \Pr\left(F_{t+1}|\Im_{t+1}\right) \tag{50}$$

It is clear from this expression that the use of the two quarters dating rule means that $Pr\left(S_t = 1|\Im_{t+1}\right)$ is a function not only of x_t but also of x_{t+1} and the entire past history of x_t. In econometric models $Pr\left(S_t = 1|\Im_t\right)$ is the basis of a likelihood and this will be

$$Pr\left(S_t = 1|\Im_t\right) = E\left[Pr\left(S_t = 1|\Im_{t+1}\right)|\Im_t\right]. \tag{51}$$

It is evident from (48), (49) and (50) that $Pr\left(S_t = 1|\Im_t\right)$ is a function of $E\left(\psi_{t+1}|\Im_t\right)$ as well as $\{\psi_{t-i}\}_{i=0}^{\infty}$. Thus, for the two-quarters rule, $Pr\left(S_t = 1|\Im_t\right)$ will not have a single index form, except in the special case where $E\left(\psi_{t+1}|\Im_t\right)$ is a function of the index $a+bx_t$. Only if the dating rule had been the "calculus" one would $Pr(S_t = 1|\Im_{t+1}) = (1-\psi_t)$ be a function of x_t only. Moreover,

the mapping between S_t and x_t will not be that from the CDF of a standard normal, as assumed in Probit models. Clearly the lesson of this analysis is that one cannot assume either the form of $\Pr(S_t = 1|\Im_t)$ or that it depends on only a contemporaneous variable x_t; it is necessary that one know how the S_t were generated in order to be able to write down the correct likelihood.

References

Bierens, H.J. (1983), "Uniform Consistency of Kernel Estimators of a Regression Function Under Generalized Conditions", *Journal of the American Statistical Association*, 77, 699-707.

Bordo, M.D. and D.C. Wheelock (2006), "When Do Stock Market Booms Occur? The Macroeconomic and Policy Environments of 20th Century Booms", *Federal Reserve Bank of St Louis Working Paper 2006-051A*

Bry, G., Boschan, C., (1971), *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*, New York, NBER.

Cashin, P. and C.J. McDermott (2002), "Riding on the Sheep's Back: Examining Australia's Dependence on Wool Exports", *Economic Record*, 78, 249-263.

Claessens, S,. M. A. Kose and M. E. Terrones (2008), "What Happens During Recessions: Crunches and Busts", *International Monetary Fund Working Paper 08/274*,

de Jong, R and T. Woutersen (2009, forthcoming), "Dynamic Time Series Binary Choice", *Econometric Theory*.

Eichengreen, B., M.W. Watson and R.S. Watson (1985), "Bank Rate Policy Under the Interwar Gold Standard: A Dynamic Probit Model", *The Economic Journal*, 95, 725-745.

Eichengreen, B., A.K. Rose and C. Wyplosz (1995), "Exchange Rate Mayhem: The Antecedents and Aftermath of Speculative Attacks", *Economic Policy*, 21, 251-312.

Estrella, A. and F.S. Mishkin (1998), "Predicting US Recessions: Financial Variables as Leading Indicators", *Review of Economics and Statistics*, LXXX, 28-61.

Hamilton, J.D., (1989), "A New Approach to the Economic Analysis of Non-Stationary Times Series and the Business Cycle", *Econometrica*, 57, 357-84.

Harding D., and A.R. Pagan, (2002), "Dissecting the Cycle: A methodological Investigation", *Journal of Monetary Economics*. 49 pages 365-381

Harding, D.. and A. R. Pagan (2006), "Synchronization of Cycles", *Journal of Econometrics*, 132, 59-79

Ibbotson, R.G., and J.J. Jaffee (1975), " 'Hot Issue' Markets", *Journal of Finance*, 30, 1027-1042.

Kauppi, H. and P. Saikkonen (2008), "Predicting U.S. Recessions with Dynamic Binary Response Models", *Review of Economics and Statistics*, 90, 777-791.

Kedem, B., (1980), *Binary Time Series*, Marcel Dekker, New York.

Q. Li and J.S. Racine (2007), *Nonparametric Econometrics* Princeton University Press, Princeton.

Lunde, A. and A.Timmermann (2004), "Duration Dependence in Stock Prices: An Analysis of Bull and Bear Markets", *Journal of Economic and Business Statistics*, 22, 253-73.

Meyn, S., (2007) *Control Techniques for Complex Networks*, Cambridge University Press.

Neftci, S.N. (1984), "Are Economic Times Series Asymmetric over the Business Cycle", *Journal of Political Economy*, 92, 307-28.

Newey W.K., and K.D., West (1987). "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix" *Econometrica* 55,703-8.

Pagan, A.R. and K. Sossounov (2003), "A Simple Framework for Analyzing Bull and Bear Markets", *Journal of Applied Econometrics*, 18, 23-46.

Racine, J.S. and Q. Li (2004), "Non-parametric Estimation of Regression Functions", *Journal of Applied Econometrics*, 119, 99-130.

Robinson, P.M. (1983), "Nonparametric Estimators for Time Series", *Journal of Time Series Analysis*, 4, 185-207.

Startz, R. (2008), "Binomial Autoregressive Moving Average Models with

an Application to U.S. Recessions" *Journal of Economic and Business Statistics*, 26, 1-8.

Watson, M.W. (1994), "Business Cycle Durations and Postwar Stabilization of the U.S. Economy", *American Economic Review*, 84, March, pp 24-46.