A Smooth Transition Logit Model of the Effects of Deregulation in the Electricity Market

A.S. Hurn, Annastiina Silvennoinen and Timo Teräsvirta

CREATES Research Paper 2014-9

Department of Economics and Business
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

A Smooth Transition Logit Model of the Effects of Deregulation in the Electricity Market

A.S. Hurn[*], **Annastiina Silvennoinen**[†]

School of Economics and Finance, Queensland University of Technology

Timo Teräsvirta[‡]

CREATES, Department of Economics and Business, Aarhus University, Denmark

and

School of Economics and Finance, Queensland University of Technology

March 2014

Abstract

The paper proposes and develops a smooth transition logit (STL) model that is designed to detect and model situations in which there is structural change in the behaviour underlying the latent index from which the binary dependent variable is constructed. The maximum likelihood estimators of the parameters of the model are derived along with their asymptotic properties and a Lagrange Multiplier test of the null hypothesis of linearity in the underlying latent index. The development of the STL model is motivated by the desire to assess the impact of deregulation in the Queensland electricity market by addressing the question of whether or not increased competition has resulted in changes in the behaviour of the spot price of electricity, specifically with respect to the well documented phenomenon of periodic abnormally high prices or price spikes. In testing this conjecture the STL model allows the timing of any change to be endogenously determined and also market participants' behaviour to change gradually over time. The main results reported in the paper provide clear evidence in support of the structural change in nature and duration of price spikes in Queensland. The endogenous dating of the structural change by the STL model agrees with the institutional detail surrounding the process of deregulation and indicates that the full effect of the policy change took about a year to occur. Notwithstanding the fact that the STL model was specifically developed to tackle a problem couched in an Australian institutional framework this research will be of general interest and applicability. In particular, it is applicable to any situation in which the impact and dating of policy changes is required and where the outcome of the policy is naturally measurable as a binary variable.

[*]e-mail: s.hurn@qut.edu.au

[†]e-mail: silvennoinen@qut.edu.au

[‡]e-mail: tterasvirta@econ.au.dk

1 Introduction

Linear models have a long and successful history in econometrics and applied macroeconomics. However, the situations in which the underlying economic theory or mere observations of a realised process support the notion of nonlinear behaviour are not uncommon and thus require the use of nonlinear models. Popular examples of these include models that are piecewise or locally linear such as switching regression models, smooth transition regression (STR) models or Markov-switching regression models. In the former models the variable determining the (linear) state is observable, whereas in Markov-switching models it is latent. In swiching regression and Markov switching models the number of states is finite, most often two or three. An STR model contains a continuum of linear states, and a transition from one state to another is controlled by a transition variable that is typically a stationary random variable but may also be deterministic such as time. Statistical modelling techniques covering specification, estimation and evaluation of STR models have been discussed in various works; for surveys see Teräsvirta (1998) and Teräsvirta (2004), for thorough treatment with examples, see Teräsvirta et al. (2010); with the focus on forecasting with nonlinear models, see Teräsvirta (2006), and for vector models, see Hubrich and Teräsvirta (2013).

A common feature of existing smooth transition models is that the variable or variables to be modelled are continuous stationary random variables. This paper introduces and develops a smooth transition logit (STL) model to model fundamental structural changes in the behaviour of the latent index underlying a binary dependent variable. The maximum likelihood estimators of the parameters of the model are derived along with their asymptotic properties. The focus is on the (more complicated) case in which the transition variable is deterministic, i.e., a time trend. In the STL model, the underlying latent index takes the familiar smooth-transition form and is therefore a nonlinear function which allows for a gradual change in the way various factors affect the probability of the event defined by the binary dependent variable. The STL model proves a natural framework in which the null hypothesis of linearity in the underlying latent index can be be tested against the alternative hypothesis of the presence of nonlinear behaviour. The paper develops a Lagrange Multiplier (LM) or score type test for this purpose. The LM approach has the benefit of requiring estimation of the linear model only, thus avoiding unnecessary and statistically untenable estimation of an STL model when the linear logit model adequately describes the data.

The development of the STL model is motivated by the desire to assess the impact of deregulation in the Queensland electricity market. Queensland is a region of the Australian National Electricity Market (NEM), which is the worlds largest interconnected power system encompassing New South Wales, Victoria, Queensland, South Australia and Tasmania. Wholesale trading in this market is conducted as a spot market where supply and demand are instantaneously matched through a centrally-coordinated dispatch process on a five-minute basis and the market clearing price is calculated for half-hourly intervals.

Starting in July 2007, the Queensland government systematically withdrew its involvement in both the generation and retailing of electricity in Queensland. The natural question then arises as to whether or not this change in policy has had a significant effect on the behaviour of market participants and, in particular, if competition in the electricity market resulted in increased strategic behaviour which in turn influenced electricity prices. To answer this question, the time-series properties of abnormally high spot prices for electricity, or "price spikes", defined as instances in which the spot electricity price exceeds a given threshold, are examined. Price spikes are a well documented feature of both the Queensland market and wholesale electricity markets more generally (Barlow, 2002; de Jong and Huisman, 2003; Escribano, *et al.*, 2002; Lucia and Schwartz, 2002; Burger *et al.*, 2003; Byström, 2005; Cartea and Figueroa, 2005; Mount et al., 2006; Kanamura and Ōhashi, 2007). The main conjecture of the paper is that the price spikes of longer duration than one-half hour interval, which is the time period over which the market clearing price is computed, have became less frequent after market deregulation with a concomitant increase in the probability of observing price spikes of half hour durations. This outcome reflects a balance between strategic behaviour by generators aimed at manufacturing higher prices and the increased investment in significant generation capacity by electricity retailers. The increased probability of price spikes lasting only one half-hour period derives from the technological advantage of the generators to manipulated their load within relatively tight timeframes over the slightly longer period required by retailers to bring their own capacity online in order to mitigate the effects of any price increases.

The primary objective of the paper then, is simply to ascertain whether the nature of the spiking process has changed in such a way that provides tangible evidence of increased strategic manipulation of prices in the electricity market. Furthermore, if there is evidence of a change in pattern of price spikes can this change be dated to coincide with moves towards increased

2

competition. It is to accomplish this objective that the STL model and the associated LM test are developed because the model allows the timing of change to be endogenously determined and for the market participants' behaviour to change gradually over time as the market participants learn the rules of engagement in the new competitive environment. It is important to appreciate at the outset that modelling and predicting the actual durations of price spikes, while potentially important objectives in their right, are beyond the scope of this paper.[1]

To preview the main results very briefly, clear evidence is found to support the presence of structural change in the latent index underlying the generation of price spiking process in Queensland. The LM test clearly rejects the null hypothesis of linearity in favour of the alternative hypothesis that the index has a nonlinear functional form. The estimation results strongly support the hypothesis that the probability of single-period price spikes has indeed increased since the introduction of FRC. The beginning of the change is found to occur around August 2008, which indicates a fairly long gestation period after the start of the deregulation process, and the full effect of the policy change on the behaviour of market participants appears to be completed by July 2009. During these twelve months, the probability of a price spike lasting for longer than a single half-hour period gradually decreases.

The rest paper is organized as follows: Section 2 introduces the smooth transition logit model and develops the estimation theory. Section 3 provides the details of a test for linearity based on the smooth transition model and contains simulation results on the small-sample behaviour of the test statistic. Section 4 discusses the theory for abnormal price events and the change in market structure. Details of the data set are laid out in Section 5, and Section 6 provides the results from the empirical analysis. Section 7 concludes.

2   The Smooth Transition Logit Model

It is assumed that the latent index y_t^* is related to factors $\boldsymbol{x}_t$ in a nonlinear fashion. For reasons explained above, the impact of the factors may have changed over time. Therefore, the smooth transition specification is chosen to describe the relationship and how it has evolved over time. This approach provides a great deal of flexibility around the manner and timing the change in

---

[1] Price spikes are particularly injurious to electricity retailers (who buy electricity at the spot price and sell it to final consumers at a heavily regulated price) and consequently forecasting their occurrence has been the subject of a fair amount of econometric investigation (Becker *et al.*, 2007; Christensen *et al.*, 2009; Christensen *et al.*, 2012; Clements *et al.*, 2013; Eichler *et al.*, 2013). Modelling the actual duration of abnormal price episodes is still in its infancy.

the market behaviour may have taken place.

Specifically,

$$y_t^* = \boldsymbol{x}_t'\boldsymbol{\phi} + \boldsymbol{x}_t'\boldsymbol{\psi}G(s_t; \gamma, c) + e_t, \tag{1}$$

where $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$ are $n \times 1$ parameter vectors, e_t are iid disturbances from the logistic distribution, and

$$G(s_t; \gamma, c) = (1 + \exp\{-\gamma(s_t - c)\})^{-1} = G(s_t), \quad \gamma > 0 \tag{2}$$

is a transition function, bounded between 0 and 1, and centered around the location parameter c. For identification purposes, the slope parameter γ is assumed positive, implying $G(s_t)$ is strictly increasing, and thereby the loading of the factors varies from $\boldsymbol{\phi}$ to $\boldsymbol{\phi} + \boldsymbol{\psi}$ as the values of s_t increase. The role of the parameter γ is to control the speed of the transition from one extreme state to another. As the value of γ increases the transition becomes rapid, and ultimately approaches a step function as $\gamma \to \infty$. At the other extreme, $\gamma = 0$ implies $G(s_t) = 1/2$, and the model reduces to the standard linear logit model.

For estimation purposes, we follow Goodwin et al. (2011) and replace γ by e^{η} in (2). The slope parameter to be estimated is then $\eta \in (-\infty, \infty)$. On one hand, the transformation avoids the need of a positivity restriction on γ. On the other, the sensitivity of the value of the likelihood function for variation in the value of γ is inversely related to the magnitude of γ. In practice this means that when γ is large, the number of observations around the transition is reduced, resulting in increased uncertainty about the exact shape of the transition and the precision of the estimate. The exponential transformation focusses the search on a much smaller range of parameter values for η, while still describing the same shape of transition as γ does in the original parameterisation. Consequently, the numerical precision of the estimate of the parameter governing the shape of the transition is much improved by this transformation.

The model defined by (1) and (2) is called the Smooth Transition Logit (STL) model. It bears a passing similarity to the Smooth Transition Regression model by Bacon and Watts (1971), the Smooth Transition Autoregressive model, see Chan and Tong (1986) and Teräsvirta (1994), and the Panel Smooth Transition Regression model by González et al. (2005). The latent index y_t^* now generates the observed dichotomous 0/1 variable y_t. Whenever y_t^* is positive, we observe $y_t = 1$, otherwise we observe $y_t = 0$. This approach yields the likelihood of any given

4

observation as

$$f(y_t, \boldsymbol{x}_t, s_t; \boldsymbol{\theta}) = F(\boldsymbol{x}_t, s_t; \boldsymbol{\theta})^{y_t} \{1 - F(\boldsymbol{x}_t, s_t; \boldsymbol{\theta})\}^{1-y_t} \tag{3}$$

where $F(\boldsymbol{x}_t, s_t; \boldsymbol{\theta})$ is the logistic cumulative distribution function and $\boldsymbol{\theta} = (\boldsymbol{\phi}', \boldsymbol{\psi}', \gamma, c)'$ is a vector of model parameters.

Writing $F_t = F(\boldsymbol{x}_t, s_t; \boldsymbol{\theta})$, the log-likelihood of the STL model is

$$L_T = \sum_{t=1}^{T} \ell_t = \sum_{t=1}^{T} \{y_t \ln F_t + (1 - y_t) \ln(1 - F_t)\}.$$

Let $z_t = \boldsymbol{x}_t'\boldsymbol{\phi} + \boldsymbol{x}_t'\boldsymbol{\psi}G(s_t)$. The score becomes

$$
\begin{aligned}
\frac{\partial L_T}{\partial \boldsymbol{\theta}} &= \sum_{t=1}^{T} \frac{\partial \ell_t}{\partial z_t} \frac{\partial z_t}{\partial \boldsymbol{\theta}} = \sum_{t=1}^{T} \{y_t \frac{F_t(1 - F_t)}{F_t} - (1 - y_t) \frac{F_t(1 - F_t)}{1 - F_t}\} \frac{\partial z_t}{\partial \boldsymbol{\theta}} \\
&= \sum_{t=1}^{T} \{y_t(1 - F_t) - (1 - y_t)F_t\} \frac{\partial z_t}{\partial \boldsymbol{\theta}} \\
&= \sum_{t=1}^{T} (y_t - F_t) \frac{\partial z_t}{\partial \boldsymbol{\theta}}
\end{aligned}
\tag{4}
$$

in which

$$\frac{\partial z_t}{\partial \boldsymbol{\theta}} = \boldsymbol{w}(t/T) = (\boldsymbol{x}_t', \boldsymbol{x}_t'G(s_t), \boldsymbol{x}_t'\boldsymbol{\psi}\frac{\partial G(s_t)}{\partial \gamma}, \boldsymbol{x}_t'\boldsymbol{\psi}\frac{\partial G(s_t)}{\partial c})' = \boldsymbol{K}(s_t)\boldsymbol{x}_t \tag{5}$$

and where

$$\boldsymbol{K}'(s_t) = \left[\begin{array}{cccc} \boldsymbol{I} & G(s_t)\boldsymbol{I} & \frac{\partial G(s_t)}{\partial \gamma}\boldsymbol{\psi} & \frac{\partial G(s_t)}{\partial c}\boldsymbol{\psi} \end{array} \right]$$

$$\partial G(s_t)/\partial \gamma = G(s_t)\{1 - G(s_t)\}(s_t - c)$$

$$\partial G(s_t)/\partial c = -\gamma G(s_t)\{1 - G(s_t)\}.$$

When transformation $\gamma = e^\eta$ is used in estimation, the corresponding partial derivatives become

$$\partial G(s_t)/\partial \eta = e^\eta G(s_t)\{1 - G(s_t)\}(s_t - c)$$

$$\partial G(s_t)/\partial c = -e^\eta G(s_t)\{1 - G(s_t)\}.$$

As the conditional expectation and variance of y_t are F_t and $F_t(1 - F_t)$, respectively, it follows

5

that the information matrix

$$
\begin{aligned}
\boldsymbol{J} \;=\;& \mathsf{E}(y_t - F_t)^2 \boldsymbol{w}_t \boldsymbol{w}_t' \\[2mm]
\;=\;& \mathsf{E}F_t(1 - F_t)
\begin{bmatrix}
\boldsymbol{x}_t \boldsymbol{x}_t' & \boldsymbol{x}_t \boldsymbol{x}_t' G(s_t) & \boldsymbol{x}_t(\boldsymbol{x}_t'\boldsymbol{\psi}\frac{\partial G(s_t)}{\partial \gamma}) & \boldsymbol{x}_t(\boldsymbol{x}_t'\boldsymbol{\psi}\frac{\partial G(s_t)}{\partial c}) \\[2mm]
 & \boldsymbol{x}_t \boldsymbol{x}_t' G(s_t) & \boldsymbol{x}_t G(s_t)(\boldsymbol{x}_t'\boldsymbol{\psi}\frac{\partial G(s_t)}{\partial \gamma}) & \boldsymbol{x}_t G(s_t)(\boldsymbol{x}_t'\boldsymbol{\psi}\frac{\partial G(s_t)}{\partial c}) \\[2mm]
 & & (\boldsymbol{x}_t'\boldsymbol{\psi}\frac{\partial G(s_t)}{\partial \gamma})^2 & (\boldsymbol{x}_t'\boldsymbol{\psi})^2\frac{\partial G(s_t)}{\partial \gamma}\frac{\partial G(s_t)}{\partial c} \\[2mm]
 & & & (\boldsymbol{x}_t'\boldsymbol{\psi}\frac{\partial G(s_t)}{\partial c})^2
\end{bmatrix}. \quad (6)
\end{aligned}
$$

Consistency and asymptotic normality of the parameters in (3) can now be established along the lines of Newey and McFadden (1994). In our case, the situation is somewhat more complicated because the transition function in our STL model is time, or t/T. This requires the use of triangular array asymptotics. For an illuminating discussion of this concept, see Hillebrand et al. (2013).

In order to prove consistency of the maximum likelihood estimators of the parameters in the STL model where $s_t = t/T$, we make the following assumptions:

Assumption C1. The true parameter vector is given by $\boldsymbol{\theta}_0 = (\boldsymbol{\phi}_0', \boldsymbol{\psi}_0', \gamma_0, c_0)' \in \Theta$, which is a compact space.

Assumption C2. The slope parameter γ_0 satisfies $\gamma_0 > \underline{\gamma} > 0$, and $\boldsymbol{\phi}_0 \neq \boldsymbol{\psi}_0$ with $\boldsymbol{\psi}_0 \neq \boldsymbol{0}$.

Note that γ_0 is bounded away from infinity because the parameter space is compact.

Assumption C3: The matrix $\boldsymbol{M}(r) = \mathsf{E}F(\boldsymbol{x}_t, r; \boldsymbol{\theta})\{1 - F(\boldsymbol{x}_t, r; \boldsymbol{\theta})\}\boldsymbol{x}_t\boldsymbol{x}_t'$ exists and is positive definite for all $r \in (0, 1)$ such that $F(\boldsymbol{x}_t, r; \boldsymbol{\theta}) \neq 0$.

In what follows, $\xrightarrow{p}$ denotes convergence in probability and $\xrightarrow{D}$ convergence in distribution. We have the following theorem:

Theorem 1. *Assume that conditions C1–C3 hold. Then the maximum likelihood estimator $\widehat{\boldsymbol{\theta}}$ is consistent for $\boldsymbol{\theta}_0$, that is, $\widehat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}_0$ as $T \to \infty$.*

Proof. See the Appendix.

To prove asymptotic normality of the maximum likelihood estimator with $s_t = t/T$ in (3), we have to strengthen Assumption C1 as follows:

Assumption AN1: The true parameter vector $\boldsymbol{\theta}_0$ is an interior point of Θ.

We state the following result:

Theorem 2. *Assume that conditions C1–C3 and AN1 hold and that $\widehat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}_0$ as $T \to \infty$. Then $\sqrt{T}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{D} \mathcal{N}(\boldsymbol{0}, \boldsymbol{J}^{-1})$ where*

$$\boldsymbol{J} = \lim_{T \to \infty} (1/T) \sum_{t=1}^{T} \mathsf{E} F_t (1 - F_t) \boldsymbol{w}(t/T) \boldsymbol{w}'(t/T) = \int_0^1 \boldsymbol{K}(r) \boldsymbol{M}(r) \boldsymbol{K}'(r) \, dr.$$

and $\boldsymbol{M}(r)$ is defined in Assumption C3.

Proof. See the Appendix.

3   Testing Linearity

This section develops a Lagrange Multiplier type test of the linearity of the process. Before fitting an STL model to the data, the null hypothesis of linearity must be tested. As with all models of the smooth transition variety, if the true model is linear, then the smooth transition model is not identified and its parameters cannot be estimated consistently. For this reason, the testing problem is nonstandard in the sense that the standard χ^2-based asymptotic theory is not valid. A general discussion of this problem and how to construct linearity tests in this situation can be found in Teräsvirta et al. (2010, Chapter 5).

The identification problem is circumvented as in Luukkonen et al. (1988) by approximating the model defined by the alternative hypothesis. Consider the logistic cumulative distribution function

$$F(\boldsymbol{x}_t, s_t; \boldsymbol{\theta}) = (1 + \exp\{-(\boldsymbol{x}_t'\boldsymbol{\phi} + \boldsymbol{x}_t'\boldsymbol{\psi} G(s_t; \gamma, c))\})^{-1} = (1 + \exp\{-z_t\})^{-1}. \tag{7}$$

When $\gamma = 0$ in $G(s_t; \gamma, c)$, z_t becomes linear. We choose $\gamma = 0$ to be the linearity hypothesis in the STL model. When the hypothesis is valid, c and $\boldsymbol{\psi}$ are unidentified nuisance parameters, i.e., only $\boldsymbol{\phi} + \boldsymbol{\psi}/2$ is identified. In order to test this null hypothesis, we follow Luukkonen et al. (1988) and expand z_t into a Taylor series around $\gamma = 0$. Choosing the first-order Taylor series, one obtains

$$z_t = \boldsymbol{x}_t'\boldsymbol{\phi} + (1/2)\boldsymbol{x}_t'\boldsymbol{\psi} + (1/4)\gamma \boldsymbol{x}_t'\boldsymbol{\psi}(s_t - c) + R_1$$

where R_1 is the remainder. After merging terms and reparameterising,

$$z_t = \boldsymbol{x}_t'\boldsymbol{\theta}_1 + s_t \boldsymbol{x}_t'\boldsymbol{\theta}_2 + R_1 = z_t^A + R_1 \tag{8}$$

7

where $\boldsymbol{\theta}_2 = \boldsymbol{\theta}_2(\gamma)$ such that $\boldsymbol{\theta}_2(\gamma) = \boldsymbol{0}$ if and only if $\gamma = 0$. The new null hypothesis is $\boldsymbol{\theta}_2 = \boldsymbol{0}$, where $\boldsymbol{\theta}_2$ is an $n \times 1$ vector. One can test the null hypothesis using z_t^A in (8) instead of z_t in (7). The former is linear in parameters and, besides, $R_1 = 0$ under linearity because the order of the Taylor approximation then equals zero. This being the case, the term $\exp\{-R_1\} = 1$ and $z_t = z_t^A$, so the asymptotic inference is not affected by the remainder when the null hypothesis is valid. This is the case when the test is based on the LM principle, which only requires estimation of the model under H_0. Generalising this to higher-order Taylor series expansions is straightforward.

We have thus defined an *auxiliary logit model*, in which the cumulative distribution function equals

$$F^A(\boldsymbol{x}_t, s_t; \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = (1 + \exp\{-(\boldsymbol{x}_t'\boldsymbol{\theta}_1 + s_t\boldsymbol{x}_t'\boldsymbol{\theta}_2)\})^{-1} = (1 + \exp\{-z_t^A\})^{-1}. \qquad (9)$$

From (9) one obtains

$$\frac{\partial z_t^A}{\partial \boldsymbol{\theta}_1} = \boldsymbol{x}_t \ \text{ and } \ \frac{\partial z_t^A}{\partial \boldsymbol{\theta}_2} = s_t\boldsymbol{x}_t$$

so that the second block of the 'auxiliary score' is

$$\frac{\partial L_T(\boldsymbol{x}_t, s_t; \boldsymbol{\theta}_1, \boldsymbol{\theta}_2)}{\partial \boldsymbol{\theta}_2} = \sum_{t=1}^{T} \{y_t - F^A(\boldsymbol{x}_t, s_t; \boldsymbol{\theta}_1, \boldsymbol{\theta}_2)\} s_t\boldsymbol{x}_t.$$

The score of the auxiliary logit model estimated under H_0 becomes

$$\begin{bmatrix} \frac{\partial L_T}{\partial \boldsymbol{\theta}_1} \\ \frac{\partial L_T}{\partial \boldsymbol{\theta}_2} \end{bmatrix}\Big|_{H_0} = \sum_{t=1}^{T} \begin{bmatrix} \boldsymbol{0} \\ \{y_t - F^A(\boldsymbol{x}_t, s_t; \widetilde{\boldsymbol{\theta}}_1, \boldsymbol{0})\} s_t\boldsymbol{x}_t \end{bmatrix} \qquad (10)$$

where $\widetilde{\boldsymbol{\theta}}_1$ is the maximum likelihood estimator of $\boldsymbol{\theta}_1$ estimated under linearity.

In order to construct the test, we assume that the maximum likelihood estimators of the parameters of the linear logit model are consistent and asymptotically normal when $s_t = t/T$. In fact, this is a special case of Theorems 1 and 2. To proceed, note that the tth element of the estimated covariance matrix of the score under H_0 has the form

$$\boldsymbol{D}_t = \begin{bmatrix} \boldsymbol{D}_{11t} & \boldsymbol{D}_{12t} \\ \boldsymbol{D}_{21t} & \boldsymbol{D}_{22t} \end{bmatrix}\Big|_{H_0} = \widetilde{F}_t^A(1 - \widetilde{F}_t^A) \begin{bmatrix} \boldsymbol{x}_t\boldsymbol{x}_t' & (t/T)\boldsymbol{x}_t\boldsymbol{x}_t' \\ (t/T)\boldsymbol{x}_t\boldsymbol{x}_t' & (t/T)^2\boldsymbol{x}_t\boldsymbol{x}_t' \end{bmatrix}$$

where $\widetilde{F}_t^A = F^A(\boldsymbol{x}_t, t/T; \widetilde{\boldsymbol{\theta}}_1, \boldsymbol{0})$ is the maximum likelihood (plug-in) estimator of $F_t^A = F_t^A(\boldsymbol{x}_t, t/T; \boldsymbol{\theta}_1, \boldsymbol{0})$.

This yields the average

$$\boldsymbol{D}_T = (1/T)\sum_{t=1}^{T}\boldsymbol{D}_t = (1/T)\sum_{t=1}^{T}\widetilde{F}_t^A(1-\widetilde{F}_t^A)\begin{bmatrix} \boldsymbol{D}_{11t} & \boldsymbol{D}_{12t} \\ \boldsymbol{D}_{21t} & \boldsymbol{D}_{22t} \end{bmatrix}. \tag{11}$$

The $(2,2)$ block of the inverse of $\boldsymbol{D}_T$ (positive definiteness assumed) equals

$$\boldsymbol{F}_T = (\boldsymbol{D}_{22T} - \boldsymbol{D}_{21T}\boldsymbol{D}_{11T}^{-1}\boldsymbol{D}_{12T})^{-1} \tag{12}$$

where

$$\begin{aligned} \boldsymbol{D}_{11T} &= (1/T)\sum_{t=1}^{T}\widetilde{F}_t^A(1-\widetilde{F}_t^A)\boldsymbol{x}_t\boldsymbol{x}_t' \\ \boldsymbol{D}_{12T} &= \boldsymbol{D}_{21T} = (1/T)\sum_{t=1}^{T}\widetilde{F}_t^A(1-\widetilde{F}_t^A)(t/T)\boldsymbol{x}_t\boldsymbol{x}_t' \end{aligned}$$

and

$$\boldsymbol{D}_{22T} = (1/T)\sum_{t=1}^{T}\widetilde{F}_t^A(1-\widetilde{F}_t^A)(t/T)\boldsymbol{x}_t\boldsymbol{x}_t'.$$

The LM statistic has the following form, see for example Godfrey (1988, p. 14),

$$LM_{\mathrm{NL}} = T^{-1}\frac{\partial L_T}{\partial\boldsymbol{\theta}_2'}|_{\mathrm{H}_0}\boldsymbol{F}\frac{\partial L_T}{\partial\boldsymbol{\theta}_2}|_{\mathrm{H}_0} = T^{-1}\frac{\partial L_T}{\partial\boldsymbol{\theta}_2'}|_{\mathrm{H}_0}(\boldsymbol{D}_{22}-\boldsymbol{D}_{21}\boldsymbol{D}_{11}^{-1}\boldsymbol{D}_{12})^{-1}\frac{\partial L_T}{\partial\boldsymbol{\theta}_2}|_{\mathrm{H}_0} \tag{13}$$

where $\partial L_T/\partial\boldsymbol{\theta}_2|_{\mathrm{H}_0}$ is defined in (10) and $\boldsymbol{F} = \lim_{T\to\infty}\boldsymbol{F}_T$. We make the following assumption:

Assumption C3A: The matrix $\boldsymbol{M}^A(r) = \mathsf{E}F(\boldsymbol{x}_t,r;\boldsymbol{\theta}_1,\boldsymbol{0})\{1-F(\boldsymbol{x}_t,r;\boldsymbol{\theta}_1,\boldsymbol{0})\}\boldsymbol{x}_t\boldsymbol{x}_t'$ exists and is positive definite for all $r \in (0,1)$ such that $F(\boldsymbol{x}_t,r;\boldsymbol{\theta}_1,\boldsymbol{0}) \neq 0$.

The matrix $\boldsymbol{F}$ in (13) is obtained by using the following lemma:

Lemma 3. *Assume that the condition C3A holds. Then the covariance matrix (11) has the following limit as $T \to \infty$:*

$$\boldsymbol{D} = \lim_{T\to\infty}(1/T)\sum_{t=1}^{T}\mathsf{E}\widetilde{F}_t^A(1-\widetilde{F}_t^A)\begin{bmatrix} \boldsymbol{D}_{11t} & \boldsymbol{D}_{12t} \\ \boldsymbol{D}_{21t} & \boldsymbol{D}_{22t} \end{bmatrix} = \begin{bmatrix} \int_0^1 \boldsymbol{M}^A(r)\,dr & \int_0^1 r\boldsymbol{M}^A(r)\,dr \\ \int_0^1 r\boldsymbol{M}^A(r)\,dr & \int_0^1 r^2\boldsymbol{M}^A(r)\,dr \end{bmatrix}$$

where $\boldsymbol{F}_t^A$ is as in (9) with $\boldsymbol{\theta}_2 = \boldsymbol{0}$, and $\boldsymbol{M}^A(r)$ is defined in Assumption C3A.

Proof: See the Appendix.

Under H_0, assuming that the result in Lemma 3 holds, LM_{NL} has an asymptotic χ^2-distribution with n degrees of freedom. In practice, the covariance matrix $\boldsymbol{D}$ is replaced by its consistent estimator (11).

In order to find out how the linearity test performs in small samples, we conduct a small simulation experiment. We use a logit model with intercept and two stochastic variables x_{2t} and x_{3t} such that

$$(x_{2t}, x_{3t})' \sim \text{iid}\mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 2\rho \\ 2\rho & 4 \end{bmatrix}\right)$$

where $\rho = -0.5, 0, 0.5$. In the auxiliary logit model (8), $\boldsymbol{x}_t = (1, x_{2t}, x_{3t})'$, $\boldsymbol{\theta}_1 = (0, 1, 1)$, $\boldsymbol{\theta}_1 = (1, 1, 1)$, or $\boldsymbol{\theta}_1 = (-1, 1, 1)$, and $s_t = t/T$. We test two null hypotheses. The first one is $\boldsymbol{\theta}_2 = \boldsymbol{0}$. Since in the application the null hypothesis is restricted to the intercept, we also test the hypothesis $\theta_{21} = 0$ assuming that $\theta_{22} = \theta_{23} = 0$ both under the null and the alternative. We use four sample sizes, 50, 100, 200 and 500, report the outcomes for three significance levels, 0.01, 0.05 and 0.1, and carry out 10000 replications for each of the 36 designs.

T	$H_0 : \boldsymbol{\theta}_2 = \boldsymbol{0}$			$H_0 : \theta_{21} = 0$		
	1%	5%	10%	1%	5%	10%
50	0.0085	0.0542	0.1131	0.0118	0.0635	0.1242
100	0.0082	0.0497	0.1026	0.0102	0.0538	0.1044
200	0.0096	0.0514	0.1025	0.0114	0.0499	0.1031
500	0.0087	0.0525	0.1012	0.0096	0.0460	0.0955

Table 1: Size study results. The table presents the rejection frequencies at 1%, 5%, and 10% level of significance for the two null hypotheses, $\boldsymbol{\theta}_2 = \boldsymbol{0}$ and $\theta_{21} = 0$, for sample sizes 100, 200, and 500. The experiment design has $\rho = 0.5$ and $\boldsymbol{\theta}_1 = (1, 1, 1)$. The results are based on $10\,000$ replications.

It turns out that changing ρ and varying $\boldsymbol{\theta}_1$ do not have any effect on the results. For this reason, in Table 1 we only report results for one experiment, in which the null hypothesis is either $\boldsymbol{\theta}_2 = \boldsymbol{0}$ or $\theta_{21} = 0$. In all the experiments, the empirical size is very close to the nominal size and this is true for even the smallest sample size considered. It does not matter which one of the two null hypotheses is being tested: the outcome remains the same. Notwithstanding the limited scope of the simulation experiment, our contention is that the empirical size of the LM test of linearity of the latent index underlying the STL model does not suffer from any size distortions.

4 Price Spikes in Electricity Markets

Price spikes in the electricity market occur when the spot price of electricity exceeds a given price threshold, taken here to be \$80/MWh. The threshold price that is often used in the Australian context is \$100/MWh (Becker et al., 2007; Christensen *et al.*, 2009; Christensen *et al.*, 2012) but the lower figure used in this research is chosen to reflect the marginal cost of electricity generation by retailers in the Queensland region, where peaking plants are in mainly gas-fired. The standard explanation for the occurrence of abnormal price events is a micro-theoretic one. If demand rises to the point of system capacity (due perhaps to extreme weather conditions) or if a significant portion of supply suddenly goes offline due to generation failure then a price spike results. In other words, price spikes are simply a manifestation of scarcity and are not due strategic behaviour on the part of market participants.

Increasingly, however, attention is being turned to strategic bidding behaviour in a competitive market as an explanation of price spikes. Strategic bidding behaviour by generators occurs as a consequence of the market clearing arrangements in the NEM. Generators are required to bid a load at each of ten price bands, ranging from the floor price $-\$1000$/MWh to the the market price ceiling (MPC) imposed by regulation, for every half-hour interval of the next day. The hypothetical half-hourly supply is therefore the sum of bid quantities in each of the bid price bands. This gives the incorrect impression that the market operates only in half-hour intervals. In fact, bid quantities can be changed (rebidding) as little as 5 minutes before actual dispatch and in these 5 minute intervals generators are dispatched from the lowest to the highest bid prices until demand is met at a price which is known as the *dispatch price*. The *spot price* of electricity paid by electricity retailers is, however, settled on an half-hourly basis and is calculated as the average of the 5 minute dispatch prices over the relevant half hour.

It is now apparent that within each half hour interval there is an incentive to behave strategically because if a spike occurs in the dispatch price for any 5 minute interval, then the average (spot) price received over the half hour will be higher than in the absence of the spike. One possible scenario is for base load generators to bid a small capacity at the lowest price band and then bid all remaining capacity at the highest price allowed (bid splitting). If the lowest bids from base load generators are fully dispatched and load is still required, the market operator must dispatch generators who have bid at higher prices. Once the price is forced up, the base load generators are able to rebid all their available capacity in the next and subsequent 5

minute intervals at the market floor price knowing that the half-hour average will be above their marginal cost due to the spike in the dispatch price.
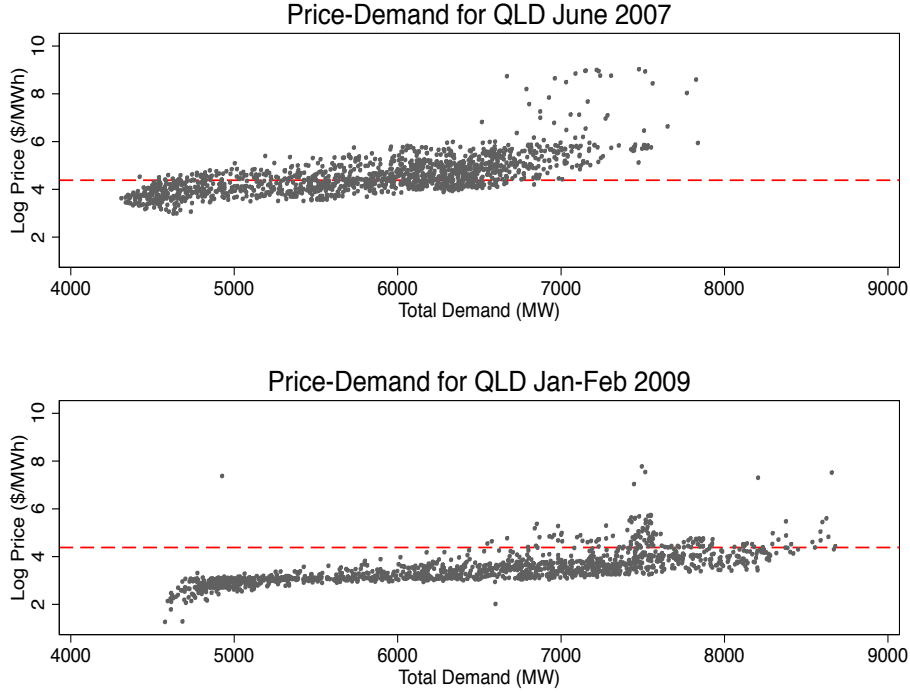

Figure 1: Scatter plots of system load versus the logarithm of price for Queensland in the months of June 2007 (top panel) and January and February 2009 (bottom panel). The dotted line represents the threshold value (natural logarithm of $80/MWh) above which an irregular price event occurs.

Figure 1, which is a scatter plot of load against the logarithm of the spot electricity price, is an illustration of the differences that can arise in the pattern of price spikes. In June 2007 Queensland was in the grip of a severe drought which limited the cooling water available for coal- and gas-fired base-load generators and also reduced the amount of water available for hydro generation. The top panel of Figure 1 therefore shows a typical situation in which the system is capacity constrained due to generation failure. Here price is regularly above the threshold price at all levels of load. The bottom panel of Figure 1, the months of January and February 2009, tells a different story with price spikes happening mainly at high loads. The fact that these extreme price events only occur at high loads, but not only at system capacity, suggests that many of these irregular events may be attributable to strategic behaviour on the part of the generators.

Starting in June 2007, the operation of the NEM in Queensland changed significantly with

the sale of two partly government-owned energy retailers and further reductions in government-owned generation capacity.[2]. In another structural development of note, privately-owned electricity retailers have embarked on a substantial program of investment in generation capacity mainly in the form of gas-fired turbines. These plants have a higher marginal cost than the base-load (mainly coal-fired) generators and therefore are used primarily as additional (peaking) capacity if conditions demand it. The fundamental point of interest for this paper is whether or not these developments, which have heralded the advent of competition in the electricity market, have had the unintended consequence of promoting strategic behaviour.

In order to test this conjecture, this paper chooses to focus on the duration of price spikes, where the duration of price spike is defined as the number of *consecutive* half hour periods that price remains above the threshold. The testable hypothesis is whether or not the durations of distinct price spikes have decreased since 2007 and is derived from the idea that strategically manufactured price spikes have a shorter duration than the episodes which are simply related to capacity shortages. If generators are able to manufacture short-lived abnormal price events then they benefit in terms of higher average prices and the retailers are unable to use their peaking generation capacity before the price spike dissipates.

On the assumption that the durations of price spike episodes, d_t, measured in half hours are available, a logistic model of these durations may be set up by defining a binary variable that takes the value 0 if the spike is a single-period spike and 1 otherwise:

$$
y_t \;\;=\;\; \begin{cases} 0 & : & d_t = 1 \\ 1 & : & d_t > 1 \,. \end{cases}
$$

The hypothesis to be tested is that $\Pr(y_t = 0)$ is unchanged after the introduction of competition, whereas the alternative is that it is increased. The problem for inference, however, is that although the date of the commencement of moves toward the competitive market is known, the effect of policy on the electricity market will have played out slowly in the subsequent years making precise inference on the effects of the policy extremely difficult. It follows that this problem is a very natural application for the the STL model in which the time of transition to competition is treated as endogenous and allowed to evolve smoothly over time.

---

[2]Government involvement in electricity generation in 2014 stands at about 50% down from around 65% in 2007 and there is no government presence in retailing.

5   Data

The data set comprises a set of durations of price spike episodes, d_t, in the Queensland spot electricity price for the period 1 January 1999 and the 31 July 2013, together with a number of control variables thought to influence this duration which are all measured at the time of the price spike. In other words, there is no time variation in the control variables. It is worth reiterating that our objective is not to model the actual duration of the episode or to make accurate predictions, but rather to ascertain if a significant change has occurred in the probability of observing price spikes that last for only one half-hour interval. In so doing our control variables are a minimum set which cover both demand and supply influences at the time the commencement of each abnormal price episode.

The discussion of the bidding process and the potential for strategic behaviour highlights the importance of load making it potentially the most important demand side control variable. Figure 2 gives an interesting perspective on the problem. The top panel of Figure 2 illustrates the average load for each half hour of the day in Queensland. This pattern is a very regular one in which there is a steep increase in demand early morning reaching peak at about 8am and then a second late flurry of demand in the late afternoon and early evening. The regularity of the pattern suggests that perhaps it is unexpected load which is more important than load itself as a control variable. While it has been noted that strategic behaviour is associated with high loads, it is also the case the unexpectedly high loads will more than likely be associated with longer abnormal price conditions. Accordingly, unexpected load, L_t, computed as the difference between an estimate of the load from a simple autoregressive model and the actual load is used in the estimation.

The bottom panel of Figure 2 shows the number of price spikes (solid line) recorded as starting in each half hour of the day (left hand scale) and the associated average duration (dashed line) of these irregular events (right hand scale). The point of interest is that the spikes earlier in the day tend to be smaller in number but longer in duration, while in the afternoon peak when demand is at its zenith, the number of spikes increases dramatically but the duration of each irregular event is very short. This observation is suggestive of the conclusion that the peak loads of the late afternoon are more likely to be associated with strategic behaviour. Accordingly the half hour of the day on which the spike starts, H_t, is included as a control variable.

Temperature effects also play an important role in contributing to episodes of price spikes.
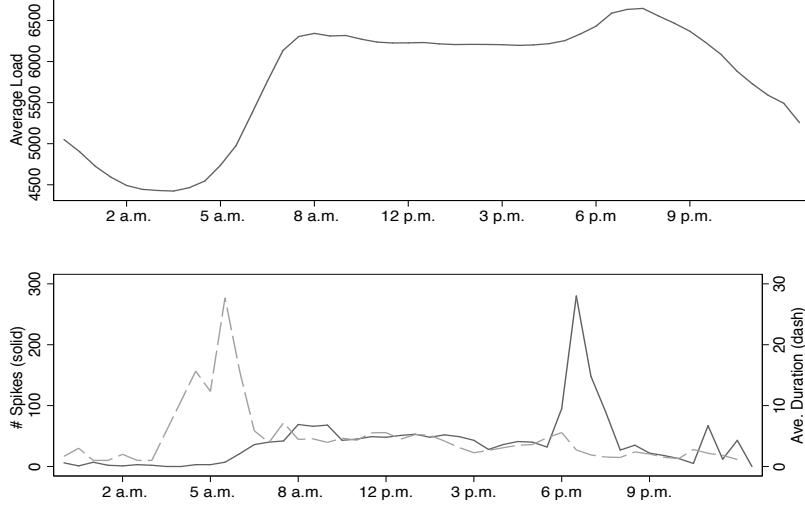
Figure 2: Plots of the average shape of load (top panel) and the number of spikes (bottom panel, solid line) and the average duration of spikes (bottom panel dashed line) by time of day.

The daily temperature range TR_t as well as the heating degree days (HDD) and cooling degree days (CDD) are also included in our modelling analysis. The cumulative heating degree days ($CH_{N,t}$) and cumulative cooling degree days ($CC_{N,t}$) indices over a period of N consecutive days are defined respectively by

$$
\begin{aligned}
CH_{N,t} &= \textstyle\sum_{k=1}^{N} HDD_{t-k}, & HDD_t &= \max\left(T - \overline{T}_t, 0\right) \\
CC_{N,t} &= \textstyle\sum_{k=1}^{N} CDD_{t-k}, & CDD_t &= \max\left(\overline{T}_t - T, 0\right)
\end{aligned}
\tag{14}
$$

where T is a threshold temperature and $\overline{T}_t$ is the average temperature on the tth day measured in degrees Celsius. The choice of threshold, in this instance $18°$C, is set by convention and is the standard used in the United States. In the southern (northern) hemisphere the HDD (CDD) season would be from May to September, while the CDD (HDD) season would be from November to March. These measures provide an indicator of system stress. If abnormally hot or cold weather persists for a number of days then it is likely that the capacity of the system will be tested.

The only connection between Queensland and the other parts of the NEM is via two regional transmission lines to the adjacent New South Wales region. The interconnector flow between the connected region of New South Wales and Queensland, QNI_t, is also used as a control variable. The convention for measuring the flow is in a northerly direction which means that if

15

QNI_t is positive, electricity is flowing into Queensland from New South Wales. The main idea being captured with this variable is that of inter-regional transmission constraints. The higher the flow to Queensland the more severe is the inter-regional capacity constraint. A positive interconnector flow indicates that the Queensland system is already under stress, which makes it more likely that a price spike episode, when it occurs, will be of longer duration.

Table 2

Summary statistics of the variables used in the estimation of the STL model of price spikes in Queensland. The data set is constructed using daily data for the period 1 January 1999 to 31 July 2013.

	d_t	y_t	L_t	H_t	TR_t	$CH_{3,t}$	$CC_{3,t}$	QNI_t
pre-July 2007								
mean	3.823	0.608	29.209	29.689	10.123	3.344	9.766	286.838
sd	5.424	0.488	90.571	9.646	3.713	5.144	9.365	402.935
max	74.000	1.000	377.000	47.000	23.000	21.400	32.950	1547.700
min	1.000	0.000	-229.000	0.000	0.600	0.000	0.000	-350.000
post-July 2007								
mean	3.195	0.557	52.173	29.221	10.065	3.636	9.641	445.242
sd	3.572	0.497	83.843	8.421	3.717	5.623	9.302	476.714
max	27.000	1.000	311.000	47.000	20.300	22.100	29.500	1535.150
min	1.000	0.000	-247.000	4.000	2.000	0.000	0.000	-5234.820
Total								
mean	3.619	0.592	36.679	29.536	10.104	3.439	9.726	338.368
sd	4.907	0.492	89.072	9.266	3.713	5.305	9.343	434.613
max	74.000	1.000	377.000	47.000	23.000	22.100	32.950	1547.700
min	1.000	0.000	-247.000	0.000	0.600	0.000	0.000	-5234.820

Table 2 provides the summary statistics of the variables used in the estimation for the periods before and after July 2007. Mean durations decrease slightly both in raw form and in 0/1 form but the significance of these figures is difficult to assess. A different story emerges if y_t is explored in terms of proportions. In the period prior to July 2007, 39% of price spike episodes lasted for one period, but in the post-July 2007 period this proportion rose to 44%. A simple t-test of difference in proportions yields a p-value of 0.0206 of the null hypothesis of equal proportions. So there is strong evidence to suggest that there is a change in the nature of the spiking process which we would argue is consistent with increased strategic behaviour. Interestingly enough, the 6% increase in the number of single-period abnormal events is observed despite the fact that average unexpected load and interconnector flow both increased, both of which would support price spike episodes of longer duration.

6 Estimation Results

As argued previously, our interest lies in ascertaining whether or not the probability of a single-period price spike has changed significantly after the deregulation of the Queensland electricity market, while controlling for the market conditions at the times of the price events. Therefore, we are focusing on the constant term in the STL model, and testing its constancy against a smooth transition over time. Linearity is clearly rejected (value of the LM type statistic with one degree of freedom is 17.2 with p-value less than 10^{-4}). This lends strong support to our hypothesis that the strategic behaviour may have changed. To examine the quality of such change, we set out to estimate a range of nonlinear models, thus hoping to ensure robust and reliable conclusions.

Table 3 reports the results of simple logit models of the duration of irregular price events in the Queensland electricity markets. The linear logit model serves as a crude overall description of the role of the various factors, and offers no opportunity to model any structural change in the electricity market. It useful in ascertaining that the variables proposed as control factors do indeed have significant explanatory power in terms of the probability of observing an irregular price episode of longer than half hour period. The effect of the unexpected load on the price spike duration is as expected. The higher the unexpected load, the higher the probability of a price spike of longer than half an hour. The time of the day variable has a strongly negative effect on the duration. This picks up the effect observed in Figure 2 that one-period spikes predominantly occur later in the day. An interesting phenomenon is that the daily temperature range has a negative effect on price spike duration. Temperature range exhibits a cyclical pattern such that a high ranges coincides with the winter months whereas during summer the difference in daily maximum and minimum temperature is small. The estimated effect may be interpreted as picking up the fact that the summer months experience much higher loads and the system operates much closer to capacity. The is no doubt that the system operates under higher stress in the Queensland summer months leading to a higher probability of longer price spikes. Conversely, the loads in winter are much lower and many generators experience scheduled maintenance so that the expectation is for shorter abnormal price events, all other things equal. The cumulative heating and cooling day variables are both positive indicating that persistently hot or cold weather will contribute to system stress and make the duration of irregular price events longer. Finally, the higher the interconnector flow from New South Wales

17

Table 3

Estimated linear logit model, regime switching logit regressions with dummy variables that determine the onset of full retail competition at different points in time, and the smooth transition logit model with estimated location and speed of transition.

	linear logit	RS–logit (2007)	RS-logit (2008)	RS-logit (2009)	STL
L_t	0.0010 (0.0005)	0.0012 (0.0005)	0.0012 (0.0005)	0.0012 (0.0005)	0.0013 (0.0005)
H_t	−0.0151 (0.0048)	−0.0163 (0.0049)	−0.0167 (0.0049)	−0.0166 (0.0049)	−0.0167 (0.0049)
TR_t	−0.0426 (0.0139)	−0.0454 (0.0140)	−0.0483 (0.0141)	−0.0475 (0.0141)	−0.0480 (0.0141)
$CH_{3,t}$	0.0532 (0.0117)	0.0551 (0.0117)	0.0600 (0.0119)	0.0592 (0.0119)	0.0597 (0.0119)
$CC_{3,t}$	0.0184 (0.0066)	0.0197 (0.0066)	0.0238 (0.0067)	0.0220 (0.0067)	0.0238 (0.0068)
QNI_t	0.0008 (0.0001)	0.0009 (0.0001)	0.0009 (0.0001)	0.0009 (0.0001)	0.0009 (0.0001)
Constant ϕ	0.5952 (0.2351)	0.7391 (0.2389)	0.7119 (0.2384)	0.7164 (0.2383)	0.7173 (0.2392)
$D_t^{\text{Jan 2007}}$		−0.4177 (0.0958)			
$D_t^{\text{Jan 2008}}$			−0.6741 (0.1127)		
$D_t^{\text{Jan 2009}}$				−0.7730 (0.1256)	
Constant ψ					−0.8055 (0.1252)
η					5.0199 (1.6792)
c					0.7162 (0.0103)
N	2284	2284	2284	2284	2284
log-likelihood	−1498.845	−1489.306	−1480.840	−1479.744	−1476.998
BIC	3051.826	3040.481	3023.549	3021.357	3031.333

Standard errors in parentheses

into Queensland, indicating the Queensland system is already struggling to cope with demand, the higher the probability of a price spike with duration longer than half an hour.

We now turn to the question of particular interest, namely, whether or not there is a change in the probability of observing one-period price spikes associated with the move to full retail competition. Accordingly, three regime switching (RS–logit) models are estimated, each of which include a dummy variable to account for structural change. These dummy variables are defined

as

$$D_t^j = \begin{cases} 1 & : & t \geq j \quad j = \text{Jul } 2007, \text{Jul } 2008, \text{Jul } 2009 \\ 0 & : & t < j \,. \end{cases}$$

Qualitatively, we can conclude that the last three models offer a better fit in terms of likelihood value than the first two. Another observation is that while the moves to full competition in the Queensland market started in mid-2007, the models with a structural change dated a year or two later seem to fit the data better than the model in which the threshold is dated on July 2007. The estimated values of the coefficients on the control variables remain practically unchanged. However, the probability of a single-period price event is significantly reduced when compared to the linear logit model, and the magnitude of reduction is greater the later the break in structure is imposed. This is in line with the presumption that the strategic behaviour may have taken a while to adjust to the policy change.

The next step is to estimate the STL model. The last column of Table 3 reports the results. As with the regime switching models, the resulting estimates, significance and sign, on the various control variables remain unchanged. The fit of the model is somewhat improved in comparison to the models just discussed, thus providing evidence that an instantaneous change in probability of long versus short price events does not enjoy as much support from the data, but that a more gradual transition does. Furthermore, the increased probability of observing half-hour price spikes is most stark in the STL model. The location of the transition is centered around mid January 2009. However, the transition takes place over several months, starting August 2008 and completed July 2009. Approximately 120 price events happened during the transition, out of which about 70 were longer than half an hour in duration.

An interesting detail in Table 3 is that BIC prefers the RS-logit (2008) and (2009) models to the STL model. This raises the question of testing the null hypothesis $\gamma = \infty$ (threshold model) against the STL one ($\gamma < \infty$). The likelihood ratio (LR) test would seem an obvious choice, but it cannot be applied to this problem. The reason is that because of the threshold the log-likelihood function is not well behaved under the null hypothesis. As a result, the asymptotic null distribution of the LR statistic is not known. However, the STL model offers a more plausible interpretation than a threshold model. It is more natural to think that both the electricity generators and retailers learn their new strategic behaviour gradually rather than to presume that they do it simultaneously at a given point of time.
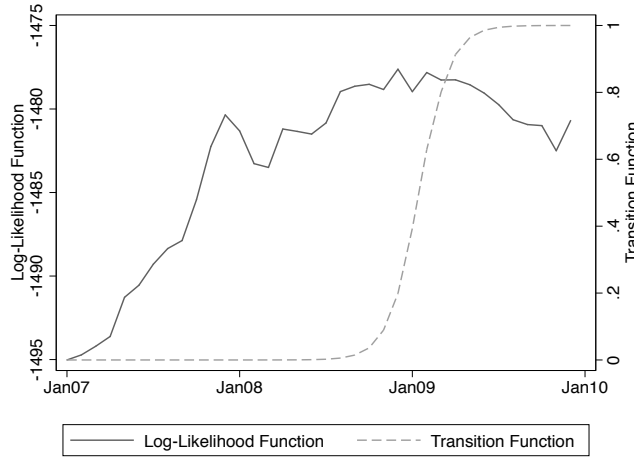
19

Figure 3: Simple grid to date the structural change in strategic behaviour by participants in the Queensland electricity market. The value of the log-likelihood function from the estimation of the logit model which includes a dummy variable dating the change at each month in the interval January 2007 to December 2009 (left axis) is plotted together with the estimated transition function from the smooth transition model (right axis).

As a check on the results of the STL model, 36 simple RS–logit models were estimated with each regression including as an explanatory variable the dummy variable defined as

$$
D_t^j \;\; = \;\; \begin{cases} 1 & : & t \geq j \quad j = \text{Jan } 2007, \text{Feb } 2007, \cdots \text{Nov } 2009, \text{Dec } 2009 \\ 0 & : & t < j \, . \end{cases}
$$

For each logit regression the value of the log-likelihood function was recorded. The results of this rather crude grid search are presented in Figure 3 (left axis) together with the estimated transition function from the smooth transition logit model based on the parameter estimates reported in Table 3. The estimated location for the smooth transition coincides with the threshold model with the highest likelihood values.

7 Conclusions

This paper introduces a smooth transition logit model for a binary response variable. The maximum likelihood estimators of the parameters of the STL model are shown to be consistent and asymptotically normal. A test of the linearity of the process generating the latent probability index underlying the binary dependent variable is also provided. Notwithstanding the fact that the STL model was specifically developed to tackle a problem couched in an Australian institutional framework, our contention is that this research will be of general interest and

applicability. In particular, it is applicable to any situation in which the impact and dating of effective policy changes is required and where the outcome of the policy is naturally measurable as a binary variable. Potential examples which spring immediately to mind are changes in individual choices in response to policy-induced incentive changes in the area of public transport, and patient outcomes in response to changes in health policy, either organisational or funding related.

In any event, the STL model is ideally suited for exploring whether or not there has been a gradual shift in the behaviour of electricity generators in the Queensland electricity market. Prior to 2007 most electricity generation was in the hands of the public sector, but subsequent developments has seen a substantial shift with privately owned generators and retailers of electricity becoming the major players. To this end, the maintained conjecture of the paper is that the change in the competitive environment has been accompanied by an increase in strategic behaviour on the part of generators who increasingly have used bidding behaviour to cause short-lived spikes in electricity prices. The testable hypothesis is that the probability of occurrence of short price spikes of half-hour duration in the spot electricity market will have increased after the move toward increased competition, if other conditions prevailing at the time of the price spike are accounted for. Another hypothesis is that the change in the durations of the irregular price episodes will have taken place gradually over a number of years as the market participants experienced the operation of the competitive market. For the same reason, the timing of the change will have occurred at a somewhat later date than the introduction of the competitive market structure.

We provide estimation results that are conclusive and robust in terms of comparison to several simple logit models in which a structural break is imposed by including a time dummy variable. Our LM type test of linearity strongly rejects the notion that the process underlying the binary variable is a linear one. The smooth transition model performs extremely well allows an interpretation of the estimated model that supports the intuition behind the research question. The estimation results suggest that the probability of occurrence of half-hourly price spikes started to increase in late 2007, was in full swing by late 2008 and was largely completed by the middle of 2009.

References

Bacon DW, Watts DG. 1971. Estimating the transition between two intersecting straight lines. *Biometrika* **58**: 525–534.

Barlow M. 2002. A diffusion model for electricity prices. *Mathematical Finance* **12**: 287 – 298.

Becker R, Hurn A, Pavlov V. 2007. Modelling spikes in electricity prices. *Economic Record* **83**: 371 – 382.

Burger M, Klar B, Mueller A, Schindlmayr G. 2003. A spot market model for pricing derivatives in electricity markets. *Quantitative Finance* **4**: 109 – 122.

Byström HNE. 2005. Extreme value theory and extremely large electricity price changes. *International Review of Economics and Finance* **14**: 41 – 55.

Cartea A, Figueroa MG. 2005. Pricing in electricity markets: A mean reverting jump diffusion model with seasonality. *Applied Mathematical Finance* **12**: 313 – 335.

Chan KS, Tong H. 1986. On estimating thresholds in autoregressive models. *Journal of Time Series Analysis* **7**: 178–190.

Christensen T, Hurn A, Lindsay K. 2009. It never rains but it pours: Modelling the persistence of spikes in electricity prices. *Energy Journal* **30**: 25 – 48.

Christensen T, Hurn A, Lindsay K. 2012. Forecasting spikes in electricity prices. *International Journal of Forecasting* **28**: 400 – 411.

Clements A, Fuller J, Hurn A. 2013. A semi-parametric method for forecasting spikes in electricity prices. *Economic Record* **89**: 508–521.

de Jong C, Huisman R. 2003. Option pricing for power prices with spikes. *Energy Power Risk Management* **7**: 12 – 16.

Eichler M, Grothe O, Manner H, Türk D. 2013. Models for short-term forecasting of spike occurrences in australian electricity markets: a comparative study. Working paper.

Escribano A, Peña JI, Villaplana P. 2002. Modelling electricity prices: International evidence. Working Paper 02-27, Universidad Carlos III de Madrid.

Godfrey LG. 1988. *Misspecification tests in econometrics. The Lagrange multiplier principle and other approaches.* Cambridge: Cambridge University Press.

González A, Teräsvirta T, van Dijk D. 2005. Panel smooth transition regression models. SSE/EFI Working Paper Series in Economics and Finance.

Goodwin BK, Holt MT, Prestemon JP. 2011. North american oriented strand board markets, arbitrage activity, and market price dynamics: A smooth transition approach. *American Journal of Agricultural Economics* **93**: 993–1014.

Hillebrand E, Medeiros MC, Xu J. 2013. Asymptotic theory for regressions with smoothly changing parameters. *Journal of Time Series Econometrics* **5**: 133–162.

Hubrich K, Teräsvirta T. 2013. Thresholds and smooth transitions in vector autoregressive models. *Advances in Econometrics* **32**: 273–326.

Kanamura T, Ōhashi K. 2007. On transition probabilities of regime switching in electricity prices. *Energy Economics* **30**: 1158 – 1172.

Lucia JJ, Schwartz ES. 2002. Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of Derivatives Research* **5**: 5 – 50.

Luukkonen R, Saikkonen P, Teräsvirta T. 1988. Testing linearity against smooth transition autoregressive models. *Biometrika* **75**: 491–499.

Mount T, Ning Y, Cai X. 2006. Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters. *Energy Economics* **28**: 62 – 80.

Newey WK, McFadden DL. 1994. Large sample estimation and hypothesis testing. In Engle RF, McFadden DL (eds.) *Handbook of Econometrics*, volume 4. Elsevier Science, Amsterdam, 2111–2245.

Teräsvirta T. 1994. Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association* **89**: 208–218.

Teräsvirta T. 1998. Modeling economic relationships with smooth transition regressions. In Ullah A, Giles DEA (eds.) *Handbook of applied economic statistics.* New York: Dekker, 507–552.

Teräsvirta T. 2004. Smooth transition regression modelling. In Lütkepohl H, Krätzig M (eds.) *Applied Time Series Econometrics*. Cambridge University Press, 223–243.

Teräsvirta T. 2006. Forecasting economic variables with nonlinear models. In Elliott G, Granger CWJ, Timmermann A (eds.) *Handbook of economic forecasting*, volume 1. Amsterdam: Elsevier, 413–457.

Teräsvirta T, Tjøstheim D, Granger CWJ. 2010. *Modelling Nonlinear Economic Time Series*. Oxford: Oxford University Press.

A Proofs

A.1 Consistency

We introduce the following notation: $||\boldsymbol{x}||$ denotes the Euclidean norm of vector $\boldsymbol{x}$, $||\boldsymbol{X}||$ denotes the Hilbert-Schmidt norm of matrix $\boldsymbol{X}$, and $\overset{D}{\rightarrow}$ denotes convergence in distribution. To prove Theorem 1, we first formulate and prove three lemmas.

Lemma 4. *The array of expectations* $\mathsf{E}\boldsymbol{w}(t/T)\boldsymbol{w}'(t/T)$ *has the following positive definite limit as* $T \rightarrow \infty$:

$$
\begin{aligned}
\boldsymbol{Q} &= \lim_{T \to \infty} (1/T) \sum_{t=1}^{T} \mathsf{E}\boldsymbol{w}(t/T)\boldsymbol{w}'(t/T) \\
&= \lim_{T \to \infty} (1/T) \sum_{t=1}^{T} \boldsymbol{K}(t/T)\boldsymbol{M}\boldsymbol{K}'(t/T) = \int_0^1 \boldsymbol{K}(r)\boldsymbol{M}\boldsymbol{K}'(r)\,dr
\end{aligned}
$$

where $\boldsymbol{M} = \mathsf{E}\boldsymbol{x}_t\boldsymbol{x}_t'$, *and*

$$
\boldsymbol{K}'(r) = \left[\begin{array}{cccc} \boldsymbol{I} & G(r)\boldsymbol{I} & \frac{\partial G(r)}{\partial \gamma}\boldsymbol{\psi} & \frac{\partial G(r)}{\partial c}\boldsymbol{\psi} \end{array} \right].
$$

Proof. Write

$$
\mathsf{E}\boldsymbol{w}(t/T)\boldsymbol{w}'(t/T) = \boldsymbol{K}(t/T)\mathsf{E}\boldsymbol{x}_t\boldsymbol{x}_t'\boldsymbol{K}'(t/T) = \boldsymbol{K}(t/T)\boldsymbol{M}\boldsymbol{K}'(t/T)
$$

for $t = 1, ..., T$ and all T. Let $t = [Tr]$, and write

$$
\begin{aligned}
\boldsymbol{Q}_T &= (1/T) \sum_{t=1}^{T} \mathsf{E}\boldsymbol{w}(t/T)\boldsymbol{w}'(t/T) = (1/T) \sum_{t=1}^{T} \boldsymbol{K}([Tr]/T)\boldsymbol{M}\boldsymbol{K}'([Tr]/T) \\
&= \sum_{t=1}^{T} \int_{t/T}^{(t+1)/T} \boldsymbol{K}([Tr]/T)\boldsymbol{M}\boldsymbol{K}'([Tr]/T)\mathrm{d}r \\
&= \int_{1/T}^{(t+1)/T} \boldsymbol{K}([Tr]/T)\boldsymbol{M}\boldsymbol{K}'([Tr]/T)\mathrm{d}r \rightarrow \int_0^1 \boldsymbol{K}(r)\boldsymbol{M}\boldsymbol{K}'(r)\mathrm{d}r
\end{aligned}
$$

as $T \rightarrow \infty$. Now consider

$$
\boldsymbol{M} - \boldsymbol{M}(h) = \mathsf{E}[1 - F(\boldsymbol{x}_t, h; \boldsymbol{\theta})\{1 - F(\boldsymbol{x}_t, h; \boldsymbol{\theta})\}]\boldsymbol{x}_t\boldsymbol{x}_t'.
$$

Note that $0 \leq F(\boldsymbol{x}_t, h; \boldsymbol{\theta})\{1 - F(\boldsymbol{x}_t, h; \boldsymbol{\theta})\} \leq 1/2$ for all h. Then $\boldsymbol{M}$ is positive definite as

Assumption C3 states that $\boldsymbol{M}(h)$ is positive definite for $F(\boldsymbol{x}_t, h; \boldsymbol{\theta}) \neq 0$. ∎

Lemma 5. *When Assumptions C1–C3 are satisfied, the STL model (1) with (2) is identified.*

Proof. Since $\boldsymbol{M} = \mathsf{E}\boldsymbol{x}\boldsymbol{x}'$ is positive definite, then $(\boldsymbol{\theta} - \boldsymbol{\theta}_0)'\boldsymbol{K}(t/T)\boldsymbol{M}\boldsymbol{K}'(t/T)(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$ is positive definite for $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$. It follows that $\mathsf{E}\{(\boldsymbol{\theta} - \boldsymbol{\theta}_0)'\boldsymbol{K}(t/T)\boldsymbol{x}\}^2 > 0$ for $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$. Then

$$(\boldsymbol{\phi} - \boldsymbol{\phi}_0)'\boldsymbol{x} + k_1(t/T)(\boldsymbol{\psi} - \boldsymbol{\psi}_0)'\boldsymbol{x} + k_2(t/T)(\gamma - \gamma_0)\boldsymbol{\psi}'\boldsymbol{x} + k_3(t/T)(c - c_0)\boldsymbol{\psi}'\boldsymbol{x} \neq 0$$

when Assumption C2 holds, where $k_1(t/T)$, $k_2(t/T)$, and $k_3(t/T)$ are bounded nonlinear nonzero functions of t/T, $t = 1, ..., T$. Thus, $\boldsymbol{\phi}'\boldsymbol{x} \neq \boldsymbol{\phi}_0'\boldsymbol{x}$ for $\boldsymbol{\phi} \neq \boldsymbol{\phi}_0$, $\boldsymbol{\psi}'\boldsymbol{x} \neq \boldsymbol{\psi}_0'\boldsymbol{x}$ for $\boldsymbol{\psi} \neq \boldsymbol{\psi}_0$, $\gamma\boldsymbol{\psi}'\boldsymbol{x} \neq \gamma_0\boldsymbol{\psi}'\boldsymbol{x}$ for $\gamma \neq \gamma_0$ and $c\boldsymbol{\psi}'\boldsymbol{x} \neq c_0\boldsymbol{\psi}'\boldsymbol{x}$ for $c \neq c_0$. Following Newey and McFadden (1994), we argue that as the logistic functions $F(z)$ and $1 - F(z)$ are strictly monotonic, $F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta}) \neq F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta}_0)$ and $1 - F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta}) \neq 1 - F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta}_0)$. From this it follows that

$$f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta}) = F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta})^{y_t}(1 - F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta}))^{1-y_t} \neq f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta}_0)$$

for $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, so the model is identified. ∎

Lemma 6. *Under Assumptions C1–C3, $\mathsf{E}|\ln f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta})| < \infty$.*

Proof. It is seen from (1) and (2) that (3) is continuous at each $\boldsymbol{\theta} \in \Theta$ with probability 1. We have

$$
\begin{aligned}
|\ln f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta})| &= |y_t \ln F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta}) + (1 - y_t)\ln\{1 - F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta})\}| \\
&\leq |\ln F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta})| + |\ln\{1 - F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta})\}|.
\end{aligned}
$$

Using (5), we obtain the following expansion of F_t around $\overline{\boldsymbol{\theta}} \in \Theta$:

$$\ln F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta}) = \ln F(\boldsymbol{x}_t, t/T; \overline{\boldsymbol{\theta}}) + F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta}^*)\{1 - F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta}^*)\}\boldsymbol{w}^{*'}(t/T)(\boldsymbol{\theta} - \overline{\boldsymbol{\theta}})$$

where $\boldsymbol{w}^*(t/T) = \boldsymbol{K}^*(t/T)\boldsymbol{x}_t$ with

$$\boldsymbol{K}^*(t/T) = (\boldsymbol{I}, G(t/T, \gamma, c)\boldsymbol{I}, \boldsymbol{\psi}\frac{\partial G(t/T)}{\partial \gamma}|_{\gamma=\gamma^*}, \boldsymbol{\psi}\frac{\partial G(t/T)}{\partial c}|_{c=c^*})'. \tag{15}$$

26

In (15), $\gamma < \gamma^* < \overline{\gamma}$ and $c < c^* < \overline{c}$, so $\boldsymbol{\theta}^* = (\boldsymbol{\phi}', \boldsymbol{\psi}', \gamma^*, c^*)'$. We have

$$
\begin{aligned}
|\ln F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta})| &= |\ln F(\boldsymbol{x}_t, t/T; \overline{\boldsymbol{\theta}}) + \{1 - F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta}^*)\} \boldsymbol{x}_t' \boldsymbol{K}^{*\prime}(t/T)(\overline{\boldsymbol{\theta}} - \boldsymbol{\theta})| \\
&\leq |\ln F(\boldsymbol{x}_t, t/T; \overline{\boldsymbol{\theta}})| + |1 - F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta}^*)| \\
&\quad \times ||\boldsymbol{K}^{*\prime}(t/T)||^2 \times ||\boldsymbol{x}_t|| \times ||\overline{\boldsymbol{\theta}} - \boldsymbol{\theta}||.
\end{aligned}
$$

Since $\boldsymbol{\theta}, \overline{\boldsymbol{\theta}} \in \Theta$ and Θ is compact, $||\overline{\boldsymbol{\theta}} - \boldsymbol{\theta}|| \leq C_1 < \infty$. By Assumption C1 and continuity and differentiability of $G(t/T)$, the elements in $\boldsymbol{K}^*(t/T)$ are bounded, implying $||\boldsymbol{K}^{*\prime}(t/T)||^2 \leq C_2 < \infty$. Thus $|\ln F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta})| \leq |\ln F(\boldsymbol{x}_t, t/T; \overline{\boldsymbol{\theta}})| + C_1 C_2 ||\boldsymbol{x}_t||$ and, because $1 - F(z) = F(-z)$, $|\ln F(-\boldsymbol{x}_t, t/T; \boldsymbol{\theta})| \leq |\ln F(-\boldsymbol{x}_t, t/T; \overline{\boldsymbol{\theta}})| + C_1 C_2 ||\boldsymbol{x}_t||$. Thus,

$$
\begin{aligned}
|\ln f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta})| &\leq |\ln F(\boldsymbol{x}_t, t/T; \overline{\boldsymbol{\theta}})| + |\ln F(-\boldsymbol{x}_t, t/T; \overline{\boldsymbol{\theta}})| \\
&\quad + 2C_1 C_2 ||\boldsymbol{x}_t|| < \infty
\end{aligned} \tag{16}
$$

because $\mathsf{E}\boldsymbol{x}_t \boldsymbol{x}_t'$ exists, see Lemma 4. Now, (16) implies that $\mathsf{E}|\ln f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta})| < \infty$. $\blacksquare$

We are now ready to prove Theorem 1. This is done by verifying the conditions of Theorem 2.5 in Newey and McFadden (1994). They are:

(i) If $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, then $\ln f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta}) \neq \ln f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta}_0)$.

(ii) The parameter space Θ is a compact space.

(iii) $\ln f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta})$ is continuous at each $\boldsymbol{\theta} \in \Theta$ with probability 1.

(iv) $\mathsf{E}|\ln f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta})| < \infty$.

Condition (i) follows from Lemma 5 and (ii) holds because of Assumption C1. Validity of Condition (iii) follows from the structure of the logistic cumulative distribution function and $G(t/T)$. Finally, Condition (iv) is verified by Lemma 6. $\blacksquare$

Asymptotic normality

In this section we prove Theorem 2. This is done by verifying the conditions of Theorem 3.3 in Newey and McFadden (1994). They are:

The conditions of Theorem 2.5 (consistency) are valid, and

(i) $\boldsymbol{\theta}_0$ is an interior point of Θ.

(ii) $f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta})$ is twice continuously differentiable and $f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta}) > 0$ in a neighbourhood $\mathcal{N}$ of $\boldsymbol{\theta}_0$.

(iii) $\int \sup_{\boldsymbol{\theta} \in \mathcal{N}} ||\frac{\partial}{\partial \boldsymbol{\theta}} f(y, \boldsymbol{x}, t/T; \boldsymbol{\theta})|| \mathrm{d}\boldsymbol{u} < \infty$, where $\boldsymbol{u} = (y, \boldsymbol{x}')'$.

(iv) $\boldsymbol{J} = \lim_{T \to \infty}(1/T) \sum_{t=1}^{T} \mathsf{E}[\frac{\partial}{\partial \boldsymbol{\theta}} \ln f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta}_0)\{\frac{\partial}{\partial \boldsymbol{\theta}'} \ln f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta}_0)\}]$ exists and is nonsingular.

(v) $\mathsf{E} \sup_{\boldsymbol{\theta} \in \mathcal{N}} ||\frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \ln f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta})|| < \infty$.

Condition (i) follows from Assumption AN1 and (ii) holds due to the properties of the STL model. To verify (iii), recall that

$$
\begin{aligned}
f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta}) &= F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta})^{y_t}(1 - F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta}))^{1-y_t} \\
&= F_t^{y_t}(1 - F_t)^{1-y_t}. \quad (17)
\end{aligned}
$$

We need the following partial derivatives:

$$
\frac{\partial}{\partial \boldsymbol{\theta}} F_t^{y_t} = y_t F_t^{y_t-1} \frac{\partial F_t}{\partial z_t} \frac{\partial z_t}{\partial \boldsymbol{\theta}} = y_t F_t^{y_t}(1 - F_t) \frac{\partial z_t}{\partial \boldsymbol{\theta}}
$$

and

$$
\begin{aligned}
\frac{\partial}{\partial \boldsymbol{\theta}}(1 - F_t)^{1-y_t} &= (1 - y_t)(1 - F_t)^{-y_t}\left(-\frac{\partial F_t}{\partial z_t}\right) \frac{\partial z_t}{\partial \boldsymbol{\theta}} \\
&= -(1 - y_t)(1 - F_t)^{1-y_t} F_t \frac{\partial z_t}{\partial \boldsymbol{\theta}}. \quad (18)
\end{aligned}
$$

Using (17) and (18) and writing $f_t = f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta})$ we have

$$
\begin{aligned}
\frac{\partial f_t}{\partial \boldsymbol{\theta}} &= \frac{\partial}{\partial \boldsymbol{\theta}} F_t^{y_t}(1 - F_t)^{1-y_t} = \{\frac{\partial F_t^{y_t}}{\partial z_t}(1 - F_t)^{1-y_t} + F_t^{y_t} \frac{\partial}{\partial z_t}(1 - F_t)^{1-y_t}\} \frac{\partial z_t}{\partial \boldsymbol{\theta}} \\
&= \{y_t F_t^{y_t}(1 - F_t)^{2-y_t} - (1 - y_t)F_t^{y_t+1}(1 - F_t)^{1-y_t}\} \frac{\partial z_t}{\partial \boldsymbol{\theta}} \\
&= F_t^{y_t}(1 - F_t)^{1-y_t}(y_t - F_t) \frac{\partial z_t}{\partial \boldsymbol{\theta}} = f_t(y_t - F_t) \frac{\partial z_t}{\partial \boldsymbol{\theta}}. \quad (19)
\end{aligned}
$$

From (19) one obtains

$$||\frac{\partial f_t}{\partial \boldsymbol{\theta}}|| \leq |f_t(y_t - F_t)| \times ||\boldsymbol{w}(t/T)|| \leq C_1||\boldsymbol{K}(t/T)|| \times ||\boldsymbol{x}_t|| \leq C_1 C_2 ||\boldsymbol{x}_t||$$

for some positive constants C_1 and C_2 as $f_t(y_t - F_t)$ is bounded for all t and the elements of $\boldsymbol{K}(t/T)$ are bounded. Following Newey and McFadden (1994), integration over $\mathrm{d}\boldsymbol{u}$ is equivalent to summing over y (y is dichotomous) and taking an expectation over $||\boldsymbol{w}(t/T)||$. This yields

$$\int ||\frac{\partial f_t}{\partial \boldsymbol{\theta}}||\mathrm{d}\mathbf{u} \leq 2C_1 C_2 \mathsf{E}||\boldsymbol{x}_t||$$

where $\mathsf{E}||\boldsymbol{x}_t|| < \infty$ for $\boldsymbol{\theta} \in \mathcal{N}$ for $t = 1, ..., T$, and all T, because $\mathsf{E}\boldsymbol{x}_t\boldsymbol{x}_t'$ exists. Thus, (iii) holds.

(iv) Consider the gradient vector

$$\frac{\partial \ln f_t}{\partial \boldsymbol{\theta}} = (y_t - F_t)\frac{\partial z_t}{\partial \boldsymbol{\theta}}.$$

Then

$$\mathsf{E}\frac{\partial \ln f_t}{\partial \boldsymbol{\theta}}\frac{\partial \ln f_t}{\partial \boldsymbol{\theta}'} = \mathsf{E}(y_t - F_t)^2\frac{\partial z_t}{\partial \boldsymbol{\theta}}\frac{\partial z_t}{\partial \boldsymbol{\theta}'}.$$

By Lemma 6,

$$\mathsf{E}\frac{\partial \ln f_t}{\partial \boldsymbol{\theta}}\frac{\partial \ln f_t}{\partial \boldsymbol{\theta}'} = \boldsymbol{K}(t/T)\mathsf{E}F_t(1 - F_t)\boldsymbol{x}_t\boldsymbol{x}_t'\boldsymbol{K}'(t/T) = \boldsymbol{K}(t/T)\boldsymbol{M}(t/T)\boldsymbol{K}'(t/T) \tag{20}$$

where $\boldsymbol{M}(t/T) = \mathsf{E}F_t(1 - F_t)\boldsymbol{x}_t\boldsymbol{x}_t'$ is positive definite by Assumption C3, see also Lemma 4. Then

$$\boldsymbol{J}_T = (1/T)\sum_{t=1}^{T}\boldsymbol{K}(t/T)\boldsymbol{M}(t/T)\boldsymbol{K}'(t/T) \tag{21}$$

is positive definite for $T \geq T_0 > 2n + 2$. Writing $F_t = F_t(t/T)$ and $t = [Tr]$ in (21) one obtains

$$
\begin{aligned}
\boldsymbol{J}_T &= (1/T) \sum_{t=1}^{T} \boldsymbol{K}([Tr]/T)[\mathsf{E}F_t([Tr]/T)\{1 - F_t([Tr]/T)\}\boldsymbol{x}_t\boldsymbol{x}_t']\boldsymbol{K}'([Tr]/T) \\
&= (1/T) \sum_{t=1}^{T} \boldsymbol{K}([Tr]/T)[\mathsf{E}F_t([Tr]/T)\{1 - F_t([Tr]/T)\}\boldsymbol{x}_t\boldsymbol{x}_t']\boldsymbol{K}'([Tr]/T) \\
&= \sum_{t=1}^{T} \int_{t/T}^{(t+1)/T} \boldsymbol{K}([Tr]/T)[\mathsf{E}F_t([Tr]/T)\{1 - F_t([Tr]/T)\}\boldsymbol{x}_t\boldsymbol{x}_t']\boldsymbol{K}'([Tr]/T)\mathrm{d}r \\
&= \int_{1/T}^{(T+1)/T} \boldsymbol{K}([Tr]/T)[\mathsf{E}F_t([Tr]/T)\{1 - F_t([Tr]/T)\}\boldsymbol{x}_t\boldsymbol{x}_t']\boldsymbol{K}'([Tr]/T)\mathrm{d}r \\
&\to \int_{0}^{1} \boldsymbol{K}(r)\boldsymbol{M}(r)\boldsymbol{K}'(r)\mathrm{d}r = \boldsymbol{J} \quad\quad\quad\quad\quad (22)
\end{aligned}
$$

as $T \to \infty$.

(v) Consider

$$
\begin{aligned}
\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} \ln f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta}) &= (y_t - F_t)\frac{\partial^2 z_t}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} - F_t(1 - F_t)\frac{\partial z_t}{\partial\boldsymbol{\theta}}\frac{\partial z_t}{\partial\boldsymbol{\theta}'} \\
&= (y_t - F_t)\frac{\partial}{\partial\boldsymbol{\theta}}\boldsymbol{K}(t/T)\boldsymbol{x}_t - F_t(1 - F_t)\boldsymbol{K}(t/T)\boldsymbol{x}_t\boldsymbol{x}_t'\boldsymbol{K}'(t/T).
\end{aligned}
$$

Then

$$
\begin{aligned}
\mathsf{E}||\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} \ln f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta})|| &\leq \mathsf{E}\{|y_t - F_t| \times ||\frac{\partial}{\partial\boldsymbol{\theta}}\boldsymbol{K}(t/T)\boldsymbol{x}_t| \\
&\quad + |F_t(1 - F_t)| \times ||\boldsymbol{K}(t/T)\boldsymbol{x}_t\boldsymbol{x}_t'\boldsymbol{K}'(t/T)||\} \\
&\leq \mathsf{E}\{||\frac{\partial}{\partial\boldsymbol{\theta}}\boldsymbol{K}(t/T)\boldsymbol{x}_t|| + ||\boldsymbol{K}(t/T)\boldsymbol{x}_t\boldsymbol{x}_t'\boldsymbol{K}'(t/T)||\} \\
&\leq \mathsf{E}\{||\frac{\partial}{\partial\boldsymbol{\theta}}\boldsymbol{K}(t/T)|| \times ||\boldsymbol{x}_t|| + ||\boldsymbol{K}(t/T)||^2 \times ||\boldsymbol{x}_t||^2\} \\
&\leq C_3\mathsf{E}||\boldsymbol{x}_t|| + C_4\mathsf{E}||\boldsymbol{x}_t||^2 < \infty
\end{aligned}
$$

for positive constants C_3 and C_4 because (a) $|y_t - F_t|$ and $|F_t(1 - F_t)|$ are bounded by 1, (b) $||\boldsymbol{K}(t/T)||$ and $||\frac{\partial}{\partial\boldsymbol{\theta}}\boldsymbol{K}(t/T)||$ are bounded because $G(t/T)$ is continuous and infinitely many times differentiable for all $\boldsymbol{\theta} \in \Theta$ and $t \leq T$ for any T, and (c) $\mathsf{E}||\boldsymbol{x}_t||$ and $\mathsf{E}||\boldsymbol{x}_t||^2$ exist because $\mathsf{E}\boldsymbol{x}_t\boldsymbol{x}_t'$ does. ∎

A.2 The covariance matrix of the average score

In this section we give a proof of Lemma 3. Consider the covariance matrix

$$
\boldsymbol{D}_T = \begin{bmatrix} \boldsymbol{D}_{11T} & \boldsymbol{D}_{12T} \\ \boldsymbol{D}_{21T} & \boldsymbol{D}_{22T} \end{bmatrix} = (1/T) \sum_{t=1}^{T} \mathsf{E} F_t^A (1 - F_t^A) \begin{bmatrix} \boldsymbol{D}_{11t} & \boldsymbol{D}_{12t} \\ \boldsymbol{D}_{21t} & \boldsymbol{D}_{22t} \end{bmatrix} |_{\mathrm{H}_0}
$$

where

$$
\boldsymbol{D}_{11T} = (1/T) \sum_{t=1}^{T} \mathsf{E} F_t^A (1 - F_t^A) \boldsymbol{x}_t \boldsymbol{x}_t' = (1/T) \sum_{t=1}^{T} \boldsymbol{M}^A (t/T)
$$

$$
\boldsymbol{D}_{12T} = \boldsymbol{D}_{21T}' = (1/T) \sum_{t=1}^{T} (t/T) \boldsymbol{M}^A (t/T)
$$

and

$$
\boldsymbol{D}_{22T} = (1/T) \sum_{t=1}^{T} (t/T)^2 \boldsymbol{M}^A (t/T) \tag{23}
$$

with $F_t^A = F(\boldsymbol{x}_t, t/T; \boldsymbol{\theta}_1, \boldsymbol{0})$. Taking (23), writing $t/T = [Tr]/T$, where $[Tr] = t$, yields

$$
\begin{aligned}
\boldsymbol{D}_{22T} &= \sum_{t=1}^{T} \int_{t/T}^{(t+1)/T} ([Tr]/T)^2 \boldsymbol{M}^A ([Tr]/T) \mathrm{d}r = \int_{1/T}^{(T+1)/T} ([Tr]/T)^2 \boldsymbol{M}^A ([Tr]/T) \mathrm{d}r \\
&\to \int_0^1 r^2 \boldsymbol{M}^A (r) \mathrm{d}r = \boldsymbol{D}_{22}
\end{aligned}
$$

as $T \to \infty$. Likewise, $\lim_{T \to \infty} \boldsymbol{D}_{12T} = \int_0^1 r \boldsymbol{M}^A (r) \mathrm{d}r = \boldsymbol{D}_{12}$ and $\lim_{T \to \infty} \boldsymbol{D}_{11T} = \int_0^1 \boldsymbol{M}^A (r) \mathrm{d}r = \boldsymbol{D}_{11}$. ∎

*** The following result may not be needed for anything:

From Lemma 4, the fact that $\mathsf{E}\{(y_t - F_t)|\boldsymbol{x}_t\} = 0$, the law of iterated expectations and noting the limit of (20) it follows that the expected Hessian has the form

$$
\begin{aligned}
\mathsf{E} \boldsymbol{H} &= \lim_{T \to \infty} (1/T) \sum_{t=1}^{T} \mathsf{E} \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \ln f(y_t, \boldsymbol{x}_t, t/T; \boldsymbol{\theta}_0) \\
&= \lim_{T \to \infty} (1/T) \sum_{t=1}^{T} \{ \mathsf{E}(y_t - F_t) \frac{\partial^2 z_t}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} |_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} - \mathsf{E} F_t (1 - F_t) \frac{\partial z_t}{\partial \boldsymbol{\theta}} \frac{\partial z_t}{\partial \boldsymbol{\theta}'} |_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \} \\
&= - \lim_{T \to \infty} (1/T) \sum_{t=1}^{T} \mathsf{E} F_t (1 - F_t) \boldsymbol{K}(t/T) \boldsymbol{x}_t \boldsymbol{x}_t' \boldsymbol{K}'(t/T) \\
&= - \lim_{T \to \infty} (1/T) \sum_{t=1}^{T} \mathsf{E} F_t (1 - F_t) \boldsymbol{K}(t/T) \boldsymbol{M}(t/T) \boldsymbol{K}'(t/T) = -\boldsymbol{J}. ∎
\end{aligned}
$$

Research Papers
2013

CREATES
Center for Research in Econometric
Analysis of Time Series

2013-44:	Federico Carlini and Paolo Santucci de Magistris: On the identification of fractionally cointegrated VAR models with the F(d) condition
2013-45:	Peter Christoffersen, Du Du and Redouane Elkamhi: Rare Disasters and Credit Market Puzzles
2013-46:	Peter Christoffersen, Kris Jacobs, Xisong Jin and Hugues Langlois: Dynamic Diversification in Corporate Credit
2013-47:	Peter Christoffersen, Mathieu Fournier and Kris Jacobs: The Factor Structure in Equity Options
2013-48:	Peter Christoffersen, Ruslan Goyenko, Kris Jacobs and Mehdi Karoui: Illiquidity Premia in the Equity Options Market
2013-49:	Peter Christoffersen, Vihang R. Errunza, Kris Jacobs and Xisong Jin: Correlation Dynamics and International Diversification Benefits
2013-50:	Georgios Effraimidis and Christian M. Dahl: Nonparametric Estimation of Cumulative Incidence Functions for Competing Risks Data with Missing Cause of Failure
2013-51:	Mehmet Caner and Anders Bredahl Kock: Oracle Inequalities for Convex Loss Functions with Non-Linear Targets
2013-52:	Torben G. Andersen, Oleg Bondarenko, Viktor Todorov and George Tauchen: The Fine Structure of Equity-Index Option Dynamics
2014-01	Manuel Lukas and Eric Hillebrand: Bagging Weak Predictors
2014-02:	Barbara Annicchiarico, Anna Rita Bennato and Emilio Zanetti Chini: 150 Years of Italian CO2 Emissions and Economic Growth
2014-03:	Paul Catani, Timo Teräsvirta and Meiqun Yin: A Lagrange Multiplier Test for Testing the Adequacy of the Constant Conditional Correlation GARCH Model
2014-04:	Timo Teräsvirta and Yukai Yang: Linearity and Misspecification Tests for Vector Smooth Transition Regression Models
2014-05:	Kris Boudt, Sébastien Laurent, Asger Lunde and Rogier Quaedvlieg: Positive Semidefinite Integrated Covariance Estimation, Factorizations and Asynchronicity
2014-06:	Debopam Bhattacharya, Shin Kanaya and Margaret Stevens: Are University Admissions Academically Fair?
2014-07:	Markku Lanne and Jani Luoto: Noncausal Bayesian Vector Autoregression
2014-08:	Timo Teräsvirta and Yukai Yang: Specification, Estimation and Evaluation of Vector Smooth Transition Autoregressive Models with Applications
2014-09:	A.S. Hurn, Annastiina Silvennoinen and Timo Teräsvirta: A Smooth Transition Logit Model of the Effects of Deregulation in the Electricity Market