UPPSALA
UNIVERSITET

Working Paper 2021:1

Department of Statistics

The deregulation of the Queensland electricity market and a smooth transition duration model

Paulina Jonéus and Johan Lyhagen

Working Paper 2021:1
April 2021
Department of Statistics
Uppsala University
Box 513
SE-751 20 UPPSALA
SWEDEN

Title: The deregulation of the Queensland electricity market and smooth transition duration model

Author: Paulina Jonéus and Johan Lyhagen
E-mail: paulina.joneusn@statistik.uu.se; johan.lyhagen@statistik.uu.se

The deregulation of the Queensland electricity market and a smooth transition duration model

Paulina Jonéus and Johan Lyhagen*

Department of Statistics, Uppsala University

April 20, 2021

Abstract

In this paper, we introduce the smooth transition duration model, designed to model the dependence of duration on explanatory variables, allowing the duration time to vary with smooth transitions over different regimes. The proposed model is a generalization of parametric survival regression models and makes it possible to detect a non-linear behaviour when the response of interest is duration time until some event occurs. A Lagrange multiplier (LM) test of the null hypothesis of linearity is derived together with the maximum likelihood estimators of the smooth transition duration model. The practical use of the introduced model is exemplified by assessing the time between abnormal price increases in the electricity spot prices in Queensland, Australia. A deregulation might have led to a change in the behaviour of the market participants and the smooth transition duration model is used to detect and examine such possible transitions. The results show a clear support of a gradual change in the appearance of abnormal price increases.

*Paulina Jonéus is a Ph.D. student, Department of statistics, Uppsala University, Kyrkogårdsgatan 10, Box 513 (e-mail: paulina.joneus@statistics.uu.se); Johan Lyhagen is a professor, Department of statistics, Uppsala University, Kyrkogårdsgatan 10 box 513 (e-mail: johan.lyhagen@statistics.uu.se). The authors are thankful for comments by Rickard Sandberg and seminar participants at the Department of statistics, Uppsala University. A special thanks to A. Stan Hurn, Annastiina Silvennoinen and Timo Teräsvirta.

1 Introduction

Duration modeling appear and is of interest in a diversity of situations within economics. Duration data measures the time in a certain state, denoted duration time, and interest lies in modeling the time until exit from that state, often with explanatory variables. A common example is movement between labor market states of employment, for treatments of duration models see e.g. Kiefer (1988) and Van Den Berg (2001). When considering unemployment, there are factors which are not obvious how to incorporate. The unemployment spells are shorter in boom while considerably longer in times of recession, and representing the state of the business cycle as dichotomous would not be appropriate as the economy is not always in the extremes, but often somewhere between. In standard regression as well as in the time series literature, a class of models which can model this feature is the smooth transition model where there is a smooth transition between the different extreme states. In addition, interest in nonlinear models has in recent years been steadily increasing not only in cases where the underlying economic theory supports the notion of a nonlinear behaviour but also due to that nonlinear models seems to outperform linear in terms of fit and forecasting performance, see e.g. Dijk et al. (2002) and Teräsvirta et al. (2010).

In this paper, we introduce a smooth transition duration model. A continuum of states is assumed, where the transition from one regime to another is controlled for by an observed deterministic transition variable. We allow for both time-varying covariates and for the duration time to vary with smooth transitions over different regimes. The estimation of the smooth transition duration model is proposed to be done using maximum likelihood and the likelihood, scores and information matrix is given. A proof of asymptotic normality of the parameter estimates is also presented. As the model is unidentified if there is no smooth transition, it is important to test for nonlinearity and an LM-type test, in the spirit of Luukkonen et al. (1988) and Teräsvirta (1994), is developed for smooth transition durations. A Monte Carlo simulation is carried out to analyse size and size adjusted power of the proposed test.

We exemplify the use of our proposed model by an empirical application where effects of the deregulation of the energy market in South East Queensland, Australia are assessed. More specifically, we model the time between abnormal electricity spot prices. Since short price increases are suspected to be a result of a strategic bidding behaviour by electricity generators to force the spot price up, it is of great interest to examine whether this behaviour is affected by regulatory changes or not. The results show support to the existence of a gradual change in the appearance of abnormal price increases.

The rest of the paper is organized as follows. Section 2 introduces the Smooth Transition duration model, the estimation and linearity tests and Section 3 presents the empirical application. Section 4 concludes.

2 The Smooth Transition Duration model

In this section we introduce the smooth transition duration model and discuss estimation and testing. Let T be an absolutely continuous, non-negative random variable representing time until an event occurs, also denoted the duration, with distribution function $F(t) = P(T \leq t)$ and density $f(t) = dF(t)/dt$. The distribution of T can then be uniquely characterized by the hazard function, specified as

$$h(t) = \lim_{dt \to 0} \frac{P(t \leq T < t + dt) \mid T \geq t)}{dt}. \tag{1}$$

The hazard function represents the instantaneous rate at which events occur, if an event has not yet occurred at time t (Kiefer, 1988). The hazard function fully specifies the distribution and so determines the density and the survivor functions.

When modelling the duration experience for a heterogeneous population, there are explanatory variables upon which the duration time might depend. Assuming the rate to be a function of derived covariates leads to the following specification of the functional form of the hazard function

$$h(t, \mathbf{x}) = \lambda_0(t) e^{\boldsymbol{\beta}' \mathbf{x}} \tag{2}$$

where $\mathbf{x} = (\mathbf{x}_1, ..., \mathbf{x}_p)$ is a vector of independent observable variables and $\boldsymbol{\beta}$ is a vector of model parameters. The function $\lambda_0(t)$ is known as the baseline hazard. Different assumptions about the hazard can be made; in the simplest case it is assumed to be constant. A generalization allows for a power dependence on duration time, such that

$$h(t, \mathbf{x}) = \alpha t^{\alpha - 1} e^{\boldsymbol{\beta}' \mathbf{x}} \tag{3}$$

where α parameterizes the duration dependence. The hazard rate decreases with duration time if $\alpha < 1$ and increases if $\alpha > 1$.

It is of interest to allow the hazard rate to change due to, for example, an unknown threshold of a covariate or the timing of the event. Cases where a threshold is included have been studied, see among others Xiaolong and Boyett (1997) and Pons (2003) for semi-parametric approaches. Cases where the duration dependence changes at a threshold have also been evaluated, see for example Lara-Porras et al. (2005) and Castro (2013).

In this paper the duration is allowed to vary over two or more regimes specified by unknown change points depending on an observed transition variable, with smooth transitions over the regimes. This results in a hazard rate model expressed as

$$h(t, \mathbf{x}) = \alpha t^{\alpha - 1} e^{(\boldsymbol{\beta} + \boldsymbol{\psi} G(s; \mathbf{c}, \gamma))' \mathbf{x}} \tag{4}$$

where the logistic transition function $G(s; \mathbf{c}, \gamma)$ has the general form

$$G(s; \mathbf{c}, \gamma) = \left(1 + exp\left\{-\gamma \prod_{k=1}^{K} (s - c_k)\right\}\right)^{-1}, \gamma > 0 \tag{5}$$

with $c = (c_1, ..., c_K)$ being a vector of location parameters with $c_1 \leq ... \leq c_K$ and $\gamma > 0$ is controlling the slope of the function. The transition function is a bounded function of the transition variable s. The transition variable is an exogenous stationary or deterministic variable. Later, s will, due to our empirical question, be rescaled calender time, i.e. $s \in (0, 1)$.

The integrated hazard function of the smooth transition duration model is

$$\begin{aligned} H(t, \mathbf{x}) &= \int_{u=0}^{t} h(u, \mathbf{x}) du \\ &= t^{\alpha} e^{(\boldsymbol{\beta} + \boldsymbol{\psi} G(s; \mathbf{c}, \gamma))' \mathbf{x}} \end{aligned} \tag{6}$$

leading to a Weibull form with distribution function

$$\begin{aligned} F(t) &= 1 - exp\left[-H(t, \mathbf{x})\right] \\ &= 1 - exp\left[-t^{\alpha} e^{(\boldsymbol{\beta} + \boldsymbol{\psi} G(s; \mathbf{c}, \gamma))' \mathbf{x}}\right]. \end{aligned} \tag{7}$$

The conditional density function of T given the covariates can then be specified as

$$f(t) = \alpha t^{\alpha - 1} e^{(\boldsymbol{\beta} + \boldsymbol{\psi} G(s; \mathbf{c}, \gamma))' \mathbf{x}} exp\left[-t^{\alpha} e^{(\boldsymbol{\beta} + \boldsymbol{\psi} G(s; \mathbf{c}, \gamma))' \mathbf{x}}\right]. \tag{8}$$

2.1 Maximum likelihood estimation

An important generalization of the model in Equation (4) allows the covariates to be time-varying. This is in practice done by letting each duration time be viewed as multiple records, where each record correspond to an interval in which the value of the covariate is constant. This generalization imposes the need of handling both censored and truncated records.

Having a sample of N observations on durations, each with n_j records with $j = 1, ..., N$, leads to N_n observations on records. Letting $i = 1, ..., N_n$, then the vector of covariates $\mathbf{x}_i$ is assumed to be fixed for each observation. Let $c_i = 1$ if record i is an observed event and denote the observation censored, with $c_i = 0$, otherwise. The ith survival time, denoted by t_i, is only observed conditional on that it had survived up until time r_i. Also, let T_i denote the calender time corresponding to i and let T denote the total number of calender time points. As above, s_i is the transition variable at i.

Letting

$$y_i = (\boldsymbol{\beta} + \boldsymbol{\psi} G(s_i; \mathbf{c}, \gamma))' \mathbf{x}_i \tag{9}$$

for $i = 1, ..., N_n$, the log likelihood of the smooth transition duration model is

$$\begin{aligned} l(\alpha, \boldsymbol{\theta}) &= \sum_{i:c_i=1} (\ln(h_i(t_i, \mathbf{x}_i)) - H_i(t_i, \mathbf{x}_i)) - \sum_{i:c_i=0} H_i(t_i, \mathbf{x}_i) + \sum_i H_i(r_i, \mathbf{x}_i) \\ &= \sum_i c_i \ln\, \alpha + c_i(\alpha - 1) \ln\, t_i + c_i y_i - t_i^{\alpha} e^{y_i} + r_i^{\alpha} e^{y_i} \end{aligned} \tag{10}$$

where $\boldsymbol{\theta} = (\boldsymbol{\beta},\, \boldsymbol{\psi},\, \gamma,\, c)$ is of dimension $2(p+1)+2$ if $\mathbf{x}_i$ is p-dimensional.

The slope parameter γ is in the estimation replaced with e^{η}. As described in Hurn et al. (2016), the gain is twofold. As the slope parameter to be estimated is

$\eta \in (-\infty, \infty)$, the positive restriction on the slope parameter is avoided. At the same time, the uncertainty about the shape of the transition is decreased in the case of a large value of the true γ.

The score is

$$\begin{bmatrix} \frac{\partial l(\alpha,\boldsymbol{\theta})}{\partial \alpha} \\ \frac{\partial l(\alpha,\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \end{bmatrix} = \sum_i \begin{bmatrix} c_i(\alpha^{-1} + \ln(t_i)) - t_i^{\alpha} e^{y_i} \ln(t_i) + r_i^{\alpha} e^{y_i} \ln(r_i) \\ (c_i - t_i^{\alpha} e^{y_i} + r_i^{\alpha} e^{y_i}) \frac{\partial y_i}{\partial \boldsymbol{\theta}} \end{bmatrix} \tag{11}$$

where

$$\begin{aligned} \frac{\partial y_i}{\partial \boldsymbol{\theta}} &= \left(\mathbf{x}_i', \, \mathbf{x}_i' G(s_i), \, \mathbf{x}_i' \boldsymbol{\psi} \frac{\partial G(s_i)}{\partial \eta}, \, \mathbf{x}_i' \boldsymbol{\psi} \frac{\partial G(s_i)}{\partial c} \right)' \\ &= \mathbf{w}(s_i) \\ &= \mathbf{K}(s_i) \mathbf{x}_i \end{aligned}$$

and where

$$\begin{aligned} \mathbf{K}(s_i)' &= \left(\mathbf{I}, \, G(s_i), \, \frac{\partial G(s_i)}{\partial \eta} \boldsymbol{\psi}, \, \frac{\partial G(s_i)}{\partial c} \boldsymbol{\psi} \right) \\ \frac{\partial G(s_i)}{\partial \eta} &= e^{\eta} G(s_i)(1 - G(s_i))(s_i - c) \\ \frac{\partial G(s_i)}{\partial c} &= -e^{\eta} G(s_i)(1 - G(s_i)). \end{aligned}$$

The information matrix can then be specified as

$$J_i(\alpha, \boldsymbol{\theta}) = \mathbb{E} \begin{bmatrix} a_i^2 & a_i \left(c_i - t_i^{\alpha} e^{y_i} + r_i^{\alpha} e^{y_i} \right) \mathbf{w}'(s_i) \\ & \left(c_i - t_i^{\alpha} e^{y_i} + r_i^{\alpha} e^{y_i} \right)^2 \mathbf{w}(s_i) \mathbf{w}(s_i)' \end{bmatrix} \tag{12}$$

with

$$a_i = c_i(\alpha^{-1} + \ln(t_i)) - t_i^{\alpha} e^{y_i} \ln(t_i) + r_i^{\alpha} e^{y_i} \ln(r_i)$$

$$\mathbf{w}(s_i)\mathbf{w}(s_i)' = \begin{bmatrix} \mathbf{x}_i \mathbf{x}_i' & \mathbf{x}_i \mathbf{x}_i' G(s_i) & \mathbf{x}_i \mathbf{x}_i' \boldsymbol{\psi} \frac{\partial G(s_i)}{\partial \eta} & \mathbf{x}_i \mathbf{x}_i' \boldsymbol{\psi} \frac{\partial G(s_i)}{\partial c} \\ & \mathbf{x}_i \mathbf{x}_i' G(s_i)^2 & \mathbf{x}_i \mathbf{x}_i' G(s_i) \boldsymbol{\psi} \frac{\partial G(s_i)}{\partial \eta} & \mathbf{x}_i \mathbf{x}_i' G(s_i) \boldsymbol{\psi} \frac{\partial G(s_i)}{\partial c} \\ & & \left(\mathbf{x}_i' \boldsymbol{\psi} \frac{\partial G(s_i)}{\partial \eta} \right)^2 & (\mathbf{x}_i' \boldsymbol{\psi})^2 \frac{\partial G(s_i)}{\partial \eta} \frac{\partial G(s_i)}{\partial c} \\ & & & \left(\mathbf{x}_i' \boldsymbol{\psi} \frac{\partial G(s_i)}{\partial c} \right)^2 \end{bmatrix} \tag{13}$$

Assuming that the transition variable is time, as mentioned in Hurn et al. (2016), the model becomes unidentified asymptotically. As $T \to \infty$ in this case, the proportion of observations in the first regime goes to zero. The parameter vector $\boldsymbol{\beta}$, governing the first regime, as well as the parameters (γ, c) vanish from the model and, hence, becomes unidentified. Using rescaled time $s_i = t_i/T$ instead makes the model identified but triangular array asymptotics (i.e. N, T both tends to infinity) need to be used instead of standard asymptotics, see (Hillebrand et al., 2013).

Following Hurn et al. (2016), consistency and asymptotic normality of the maximum likelihood estimators can be proven with the following assumptions.

A1: The true parameter vector is given by $\boldsymbol{\theta}_0 \in \Theta$ which is a compact space

A2. The slope parameter γ_0 satisfies $\gamma_0 > 0$ and $\boldsymbol{\psi}_0 \neq \boldsymbol{\beta}_0$ with $\boldsymbol{\psi}_0 \neq \mathbf{0}$

A3. The matrix $M = \mathbb{E}\mathbf{x}\mathbf{x}'$ is positive definite and $\int e^{(\mathbf{x})p(\mathbf{x})}d\mathbf{x} < \infty$ where $p(\mathbf{x})$ is the density of $\mathbf{x}$.

A4. $\lim_{T\to\infty} N/T = p \leq p_0 < \infty$ i.e. that the ratio of price events to the total of half-hour periods is asymptotically a finite constant and that $T \to \infty$ implies $N \to \infty$

A5. The duration $t_i/T \to 0$, $i = 1, ..., N$ as $T \to \infty$

2.2 LM-type of test of linearity

The smooth transition duration model specified in (4) is linear if $\boldsymbol{\psi} = \mathbf{0}$. In that case, the parameter γ is not identified and the parameters cannot be estimated consistently. Instead, an approximation of the model, defined by the alternative hypothesis, can be used when testing non-linearity (see Luukkonen et al. (1988) and Teräsvirta (1994)). Examining the form of the logistic transition function, it is clear that if $\gamma = 0$, then $y_i = \mathbf{x}_i'\boldsymbol{\beta} + G(s_i; \gamma, c)\mathbf{x}_i'\boldsymbol{\psi}$ becomes linear. In addition, adding the assumption that $\boldsymbol{\psi} \neq \mathbf{0}$, it is linear if and only if $\gamma = 0$, and due to this fact, $\gamma = 0$ can represent the linearity hypothesis (Hurn et al., 2016).

To test the null of linearity, $y_i = f(\gamma)$ is expanded by a first order Taylor series expansion around $\gamma = 0$. That is, the function $G(s_i)$ is approximated locally around the null hypothesis and replaced in Equation (9), such that

$$\begin{aligned} f(\gamma) &= f(0) + f'(0)\gamma + R_i \\ &= \mathbf{x}_i'\boldsymbol{\beta} + \frac{\mathbf{x}_i'\boldsymbol{\psi}}{2} + \frac{\mathbf{x}_i'\boldsymbol{\psi}\gamma s_i}{4} - \frac{\mathbf{x}_i'\boldsymbol{\psi}\gamma c}{4} + R_i \\ &= \mathbf{x}_i'\left(\boldsymbol{\beta} + \frac{\boldsymbol{\psi}}{2} - \frac{\boldsymbol{\psi}\gamma c}{4}\right) + \frac{\mathbf{x}_i'\boldsymbol{\psi}\gamma s_i}{4} + R_i \\ &= \mathbf{x}_i'\boldsymbol{\phi}_1 + s_i\mathbf{x}_i'\boldsymbol{\phi}_2 + R_i \\ &= y_i^A + R_i \end{aligned} \tag{14}$$

where $\boldsymbol{\phi}_1 = \left(\boldsymbol{\beta} + \frac{\boldsymbol{\psi}}{2} - \frac{\boldsymbol{\psi}\gamma c}{4}\right)$ and $\boldsymbol{\phi}_2 = \frac{\boldsymbol{\psi}\gamma}{4}$. Since $\boldsymbol{\phi}_2 = \mathbf{0}$ if and only if $\gamma = 0$, the null hypothesis can be tested by the new null hypothesis $H_0 : \boldsymbol{\phi}_2 = \mathbf{0}$, where $\boldsymbol{\phi}_2$ is of dimension $(1 \times q)$, using the equation

$$y_i^A = \mathbf{x}_i'\boldsymbol{\phi}_1 + s_i\mathbf{x}_i'\boldsymbol{\phi}_2. \tag{15}$$

Since this is a linear hypothesis in a linear model it can be tested using standard asymptotic theory. Since the remainder $R_i = 0$ for all i under the null of linearity, the term $e^{R_i} = 1$ and the asymptotic inference is not affected by the remainder when the null hypothesis is valid. The auxiliary model used for testing linearity can now be defined as

$$\begin{aligned} h^A(t, \mathbf{x}_i) &= \alpha t^{\alpha-1} e^{y_i^A} \\ &= \alpha t^{\alpha-1} e^{\mathbf{x}_i'\boldsymbol{\phi}_1 + s_i\mathbf{x}_i'\boldsymbol{\phi}_2} \end{aligned} \tag{16}$$

It is clear that

$$\begin{aligned}\frac{\partial y_i^A}{\partial \boldsymbol{\phi}_1} &= \mathbf{x}_i' \\ \frac{\partial y_i^A}{\partial \boldsymbol{\phi}_2} &= \mathbf{x}_i' s_i\end{aligned} \tag{17}$$

and the score of the auxiliary model under H_0 is

$$\left[\begin{array}{c}\frac{\partial l(\alpha,\boldsymbol{\theta})}{\partial(\alpha,\boldsymbol{\phi}_1)'} \\ \frac{\partial l(\alpha,\boldsymbol{\theta})}{\partial \boldsymbol{\phi}_2}\end{array}\right]_{|H_0} = \sum_i \left[\begin{array}{c} 0 \\ \left(c_i - t_i^{\hat{\alpha}} e^{\hat{y}_i^A} + r_i^{\hat{\alpha}} e^{\hat{y}_i^A}\right) \mathbf{x}_i' s_i \end{array}\right] \tag{18}$$

where $\hat{\alpha}$ is the ML estimator of α and $\hat{y}_i^A$ is calculated with the ML estimator $\hat{\boldsymbol{\theta}}$ under the null hypothesis of linearity. The observed outer-product estimator of the covariance matrix is then derived. The partitioned covariance matrix can be defined as

$$\mathbf{D}_{i|H_0} = \left[\begin{array}{cc} l_{11i} & \mathbf{l}_{12i} \\ \mathbf{l}_{21i} & \mathbf{L}_{22i} \end{array}\right]_{|H_0} \tag{19}$$

where

$$\begin{aligned} l_{11i} &= \left(c_i(\hat{\alpha}^{-1} + \ln(t_i)) - t_i^{\hat{\alpha}} e^{\hat{y}_i^A} \ln(t_i) + r_i^{\hat{\alpha}} e^{\hat{y}_i^A} \ln(r_i)\right)^2 \\ \mathbf{l}_{12i} &= \left(c_i(\hat{\alpha}^{-1} + \ln(t_i)) - t_i^{\hat{\alpha}} e^{\hat{y}_i^A} \ln(t_i) + r_i^{\hat{\alpha}} e^{\hat{y}_i^A} \ln(r_i)\right) \left(c_i - t_i^{\hat{\alpha}} e^{\hat{y}_i^A} + r_i^{\hat{\alpha}} e^{\hat{y}_i^A}\right) \left[\mathbf{x}_i' \quad \mathbf{x}_i' s_i\right]' \\ &= \mathbf{l}_{21i}' \\ \mathbf{L}_{22i} &= \left(c_i - t_i^{\hat{\alpha}} e^{\hat{y}_i^A} + r_i^{\hat{\alpha}} e^{\hat{y}_i^A}\right)^2 \left[\begin{array}{cc} \mathbf{x}_i \mathbf{x}_i' & \mathbf{x}_i \mathbf{x}_i' s_i \\ & \mathbf{x}_i \mathbf{x}_i' s_i^2 \end{array}\right] \end{aligned} \tag{20}$$

where non-linearities are tested in $\mathbf{x}$. To define the LM test statistic we reformulate the covariance matrix as

$$\mathbf{D}_{i|H_0} = \left[\begin{array}{cc} \mathbf{D}_{11i} & \mathbf{D}_{12i} \\ \mathbf{D}_{21i} & \mathbf{D}_{22i} \end{array}\right]_{|H_0} \tag{21}$$

where $\mathbf{D}_{11i}$ denotes the $((p+1) \times (p+1))$ upper left corner matrix of $\mathbf{D}_{i|H_0}$, $\mathbf{D}_{12i}$ the upper right $((p+1) \times p)$ matrix where $\mathbf{D}_{12i} = \mathbf{D}_{12i}'$ and $\mathbf{D}_{22i}$ denotes the $(p \times p)$ lower right corner matrix of $\mathbf{L}_{22i}$ where p is the dimension of $\boldsymbol{\psi}$, i.e. the number of non-linear components to be tested.

Letting $\mathbf{D}_{(n)}$ denote the average

$$\mathbf{D}_{(n)} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{D}_{i|H_0}$$

and using the inverse of a partitioned matrix

$$\mathbf{F}_n = \left(\mathbf{D}_{22(n)} - \mathbf{D}_{21(n)}\mathbf{D}_{11(n)}^{-1}\mathbf{D}_{12(n)}\right)^{-1}$$

the LM statistic can be specified as

$$\begin{aligned} LM &= \frac{1}{n}\frac{\partial l\left(\alpha, \boldsymbol{\theta}\right)}{\partial \boldsymbol{\phi}_2'}_{|H_0} \mathbf{F} \frac{\partial l\left(\alpha, \boldsymbol{\theta}\right)}{\partial \boldsymbol{\phi}_2}_{|H_0} \\ &= \frac{1}{n}\frac{\partial l\left(\alpha, \boldsymbol{\theta}\right)}{\partial \boldsymbol{\phi}_2'}_{|H_0} \left(\mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}\right)^{-1} \frac{\partial l\left(\alpha, \boldsymbol{\theta}\right)}{\partial \boldsymbol{\phi}_2}_{|H_0} \end{aligned} \tag{22}$$

where $\frac{\partial l_i(\alpha,\boldsymbol{\theta})}{\partial \boldsymbol{\phi}_2'}$ is defined above, $\mathbf{D}_{ij} = \text{plim}_{n\to\infty}\mathbf{D}_{ij(n)}$ with $i, j = 1, 2$ and $\mathbf{F} = \text{plim}_{n\to\infty}\mathbf{F}_n$. Following Hurn et al. (2016), this statistic can be proved to have a χ^2 distribution with q degrees of freedom. The following additional assumption is needed:

A6. The auxiliary matrix

$$\mathbb{E}\begin{bmatrix} a_i^2 & a_i\left(c_i - t_i^{\alpha}e^{\mathbf{x}_i'\boldsymbol{\phi}_1} + r_i^{\alpha}e^{\mathbf{x}_i'\boldsymbol{\phi}_1}\right)\mathbf{w}(s_i) \\ & \left(c_i - t_i^{\alpha}e^{\mathbf{x}_i'\boldsymbol{\phi}_1} + r_i^{\alpha}e^{\mathbf{x}_i'\boldsymbol{\phi}_1}\right)^2\mathbf{w}(s_i)\mathbf{w}(s_i)' \end{bmatrix} \tag{23}$$

exists and is positive definite with the logistic transition function $G(r;\gamma,c)$ where $r \in (0,1)$.

The matrix $\mathbf{F}$ can then be found by Lemma 1 in Hurn et al. (2016). Under the null hypothesis of a linear model, assuming that the result in Lemma 1 holds, the LM-statistic has an asymptotic χ^2-distribution with p degrees of freedom.

2.2.1 Tests for multiple transitions

If the test result indicates one smooth transition, it is of interest to test for an additional one, see e.g. (Teräsvirta et al., 2010, P. 384) or Appendix A.4 in Hurn et al. (2016). This test is outlined below. In the case of two transitions, the hazard rate can be formulated as

$$h(t, \mathbf{x}) = \alpha t^{\alpha-1} e^{(\boldsymbol{\beta}+\boldsymbol{\psi}_1 G_1(s_{1,u};\mathbf{c}_1,\gamma_1)+\boldsymbol{\psi}_2 G_2(s_{2,u};\mathbf{c}_2,\gamma_2))'\mathbf{x}} \tag{24}$$

with each of the two transition functions defined as above.

A test of the null that a second transition is present is an extension of the linearity test, where the parameter γ_2 now is tested. That is, the first-order Taylor series expansion around $\gamma_2 = 0$ derived as

$$\begin{aligned} y_i &= f(\gamma_2) \\ &= \mathbf{x}_i'\boldsymbol{\beta} + \mathbf{x}_i'\boldsymbol{\psi}_1 G_1(s_{1,u};\mathbf{c}_1,\gamma_1) + \mathbf{x}_i'\boldsymbol{\psi}_2 G_2(s_{2,u};\mathbf{c}_2,\gamma_2) \end{aligned} \tag{25}$$

which yields

$$\begin{aligned} f(\gamma_2) &= f(0) + f'(0)\,\gamma_2 + R_i \\ &= \mathbf{x}_i'\boldsymbol{\phi}_1 + s_{2,u}\mathbf{x}_i'\boldsymbol{\phi}_2 + \mathbf{x}_i'\boldsymbol{\psi}_1 G_1(s_{1,u};\mathbf{c}_1,\gamma_1) + R_i \\ &= y_i^A + R_i \end{aligned} \tag{26}$$

where $\boldsymbol{\phi}_1 = \left(\boldsymbol{\beta} + \frac{\psi_2}{2} - \frac{\psi_2\gamma_2 c_2}{4}\right)$ and $\boldsymbol{\phi}_2 = \frac{\psi_2\gamma_2}{4}$. Since $\boldsymbol{\phi}_2 = \mathbf{0}$ if and only if $\gamma_2 = 0$, the null hypothesis of only one transition can be tested by the new null hypothesis $H_0 : \boldsymbol{\phi}_2 = \mathbf{0}$. As before, the ith element of the estimated covariance matrix of the score under the null can be written as

$$\mathbf{D}_{i|H_0} = \begin{bmatrix} l_{11i} & \mathbf{l}_{12i} \\ & \mathbf{L}_{22i} \end{bmatrix}_{|H_0} \tag{27}$$

where now

$$\tilde{\mathbf{w}}(s_i) = \left(\mathbf{x}_i',\ \mathbf{x}_i'\tilde{G}(s_i),\ \mathbf{x}_i'\frac{\partial y_i^A}{\partial \eta_1}_{|H_0},\ \mathbf{x}_i'\frac{\partial y_i^A}{\partial c_1}_{|H_0},\ s_i\mathbf{x}_i'\right)' \tag{28}$$

and

$$\begin{aligned} l_{11i} &= \left(c_i(\hat{\alpha}^{-1} + \ln(t_i)) - t_i^{\hat{\alpha}} e^{\hat{y}_i}\ln(t_i) + r_i^{\hat{\alpha}} e^{\hat{y}_i}\ln(r_i)\right)^2 \\ \mathbf{l}_{12i} &= \left(c_i(\hat{\alpha}^{-1} + \ln(t_i)) - t_i^{\hat{\alpha}} e^{\hat{y}_i}\ln(t_i) + r_i^{\hat{\alpha}} e^{\hat{y}_i}\ln(r_i)\right)\left(c_i - t_i^{\hat{\alpha}} e^{\hat{y}_i^A} + r_i^{\hat{\alpha}} e^{\hat{y}_i^A}\right)\tilde{\mathbf{w}}'(s_i) \\ \mathbf{L}_{22i} &= \left(c_i - t_i^{\hat{\alpha}} e^{\hat{y}_i^A} + r_i^{\hat{\alpha}} e^{\hat{y}_i^A}\right)^2 \tilde{\mathbf{w}}(s_i)\tilde{\mathbf{w}}'(s_i). \end{aligned} \tag{29}$$

The test is then equivalent to the one presented above.

2.2.2 Size and power considerations

Here the performance of the LM-test of linearity is examined with a simulation experiment where data is generated using a duration model with an intercept and two variables where

$$(x_{1i}, x_{2i})' = \text{i.i.d. } N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}\right). \tag{30}$$

First, a test of a non-linear adjustment involving only the intercept, that is, the null $\phi_{20} = 0$ is examined. Second, a test of whether both the intercept and impact of control variables are changing, $\boldsymbol{\phi}_2 = 0$, is examined. The results regarding the size of the test for six different sample sizes are presented in Table 1. The empirical size of the linearity test seems to be close to the nominal size, even if there are some differences for the smaller sample sizes considered. One can especially note that the empirical size is slightly worse when testing $\boldsymbol{\phi}_2 = 0$ compared to $\phi_{20} = 0$.

Table 1: Rejection frequencies under the true null of linearity, 10 000 replicates

	$\phi_{20} = 0$			$\boldsymbol{\phi}_2 = 0$		
T	1%	5%	10%	1%	5%	10%
150	0.020	0.080	0.139	0.031	0.108	0.182
300	0.016	0.068	0.124	0.019	0.076	0.145
600	0.012	0.058	0.107	0.017	0.063	0.118
1 200	0.011	0.050	0.104	0.013	0.060	0.116
2 400	0.011	0.054	0.102	0.013	0.055	0.106
4 800	0.010	0.050	0.100	0.010	0.052	0.105

Note: T is the sample size. The three columns to the left present the nominal sizes when testing a non-linear adjustment involving only the intercept. The three columns to the right present the nominal sizes when testing a non-linear adjustment involving the intercept and impact of the control variables.

The size adjusted power of the test is presented in Table 2. The parameters under the alternative are $\boldsymbol{\psi} = (0.25, 0.25, 0.25)$ and K is the number of parameters tested, only intercept or both intercept and impact of control variables. Results where $\boldsymbol{\psi} = (0.25, 0.25, 0.25)$ is scaled with 0.5 is presented in 5 in Appendix. The results are as expected with a power of the tests approaching one for large sample sizes and both the scaling and value of the slope parameter affecting the power. We can also, as for the empirical size, note that the test performs slightly better when testing only one parameter compared to when testing three. This has the empirical implication that one should not routinely test for nonlinearity using all variables but select a smaller subset, perhaps only the constant, based on some sound thematic theory.

3 Empirical application

The deregulation in the Australian electricity market began in 2007 when the Queensland government started to implement a move towards full retail competition. Energy prices in South East Queensland was deregulated in 2016. The main point of interest is if these deregulation led to increased incentives of strategic behaviour by market participants, affecting electricity prices.

The motivation to the duration model outlined above is to analyse if strategic price behaviour is more frequent after the deregulation compared to before. To do this we model the duration of what is denoted price episodes in the Queensland electricity market. A price episode is here defined as the duration of time, measured in half hours, between two price events. A price event is defined as one or multiple consecutive half hours with a spot price above an inflation adjusted price corresponding to $80 in 2001. As stated in Hurn et al. (2016), this figure is chosen to reflect the marginal cost of electricity generation by retailers in the Queensland region. This way it is possible to assess the impact of the deregulation by examining whether or not the time until the occurrence of an episode have decreased after the deregulation.

Table 2: Size adjusted power of the test, 10 000 replicates

Slope, γ	K	150	300	600	1 200	2 400	4 800
10	1	0.212	0.367	0.646	0.909	0.996	1.000
10	3	0.159	0.313	0.571	0.876	0.996	1.000
20	1	0.247	0.432	0.728	0.954	0.999	1.000
20	3	0.190	0.380	0.674	0.940	0.999	1.000
40	1	0.260	0.450	0.749	0.962	1.000	1.000
40	3	0.198	0.399	0.701	0.951	1.000	1.000

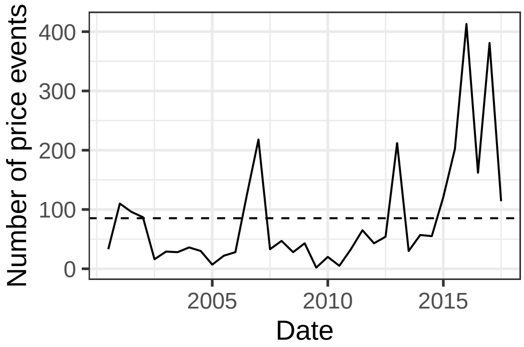
Note: The size adjusted power is presented for different sample sizes, T. The parameters under the alternative is $\boldsymbol{\psi} = (0.25, 0.25, 0.25)$ and K is the number of parameters tested; only intercept or both intercept and impact of the two control variables.

To understand the possible strategic behaviour on the electricity market, a closer look at how electricity is physically traded between generators and market consumers is needed. Six categories of participants are defined in the National Electricity Market (NEM), where the the market generators, the participants of certain interest here, sell output through the spot market and receive the spot price at settlement.

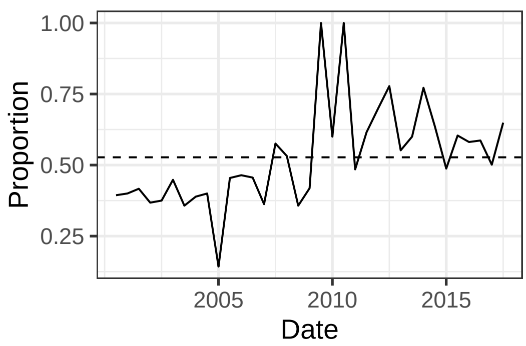
Ahead of each trading day, the market generators provide details on their availability and make an offer to produce particular quantities at particular prices. The bids indicate how many megawatts the generator wishes to produce at particular prices in up to 10 price bands. Generators can bid up to a price cap of $10,000 per MWh and the market floor price is minus $1,000 per MWh. These bids also include information about the generators minimum operating level, the point where the price bid changes from a negative to a positive price band. This is a level at which the the generator can operate indefinitely without requiring auxiliary firing to respond to changed dispatch instructions. The generators can change the bid quantities as little as 5 minutes before actual dispatch. Trading in a specific region of the NEM is then based on a 30-minute trading interval, where the spot price is the average of the six 5-minutes interval dispatch price outcomes for the preceding half hour.

Since the average spot price of an half hour interval is being inflated by an abnormally high dispatch price recorded for any 5-minute interval, one can argue that there are incentives for a strategical behaviour in the bidding process. The speed at which generating capacity can be increased is generally a lot faster than the time it takes for a generator not currently dispatched to start injecting electricity. This implies that generators already dispatched, but not on full capacity, in previous intervals are the better choice which gives them a strategic advantage. One scenario could therefore be that as the dispatched load approaches a critical point in terms of capacity, base load generators withhold capacity at the lowest price bands. If the existing bids from the 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 subsequent 5-minute-intervals. The half hour spot price is thereby forced up.

In total, the dataset comprises of half hour data from January 1, 2001 to June



(a) Number of price events



(b) Proportion of abnormal price events

Figure 1: Number of price events and proportion of all price events that lasted for exactly one half hour, per 6 months, 2001 – 2018. Dotted line represents the mean.

28, 2018, that is a total of 301 134 observations. The variable of interest is the half hour spot electricity price per MHh in Queensland, Australia. Of particular interest are the cases where the spot price is above a \$80 for one or many consecutive time points, such observations are denoted *price events*. If the price is above \$80 for exactly one half hour, the observation is denoted an *abnormal price event*.

Oppose to Hurn et al. (2016), where the aim is to find out whether the number of longer-lasting price episodes has decreased or not since 2007, we investigate another, closely related, question, namely, whether the number of consecutive half hours between two price events has decreased over time.

As mentioned in (Hurn et al., 2016), the standard explanation for the occurrence of price events might follow increased demand due to, for example, extreme weather conditions and or decreasing supply due to generation failure. Price episodes can therefore be explained by scarcity and do not, in general, imply any strategic behaviour by market participants. The abnormal price events on the other hand might be caused by incentives to increase the spot price.

There are 2 985 price events registered in the period of study, 1 558 of these are abnormal price events. The number of price events per six months are presented in Figure 1a. It can be noted that there is an increase in the number of events in the fist half of 2007, when the deregulation started, followed by a decrease. There seems to be an indication that the abnormal price events are occurring more often in recent years. The proportion of all price events that lasted for exactly one half hour was 40 percent in 2001 and increased to 63 percent in 2017, see Figure 1b. The main interest is to examine if the duration until a price event is affected by the move towards full retail competition in the electricity market, in the sense that abnormal price events might be allowed to occur more often. The time between any kind of price events is examined.

As mentioned above, temperature effects can be assumed to be important when it comes to periods of abnormal electricity spot prices. As in Hurn et al. (2016), three measures of temperature are included as control variables in the analysis of time between price events. First, the daily temperature range (TR_i) on day t_i is computed. In addition are two variables capturing extreme weather conditions constructed based on heating degree days and cooling degree days. These are computed as

Table 3: Mean and standard deviation of relevant covariates per year

Year	TR	CH	CC	QNI	UL
2001	10.58 (3.16)	2.09 (3.95)	9.56 (8.33)	-97.35 (234.74)	0.01 (79.82)
2002	10.75 (3.49)	2.38 (4.44)	10.75 (9.84)	-131.61 (287.83)	-0.04 (84.63)
2003	9.98 (3.34)	2.23 (4.05)	7.77 (7.16)	-389.26 (291.48)	0.02 (87.03)
2004	10.49 (3.78)	2.37 (4.25)	9.96 (8.98)	-397.39 (293.63)	0.03 (84.67)
2005	9.54 (2.76)	1.37 (2.73)	10.93 (8.98)	-639.51 (287.18)	-0.01 (87.98)
2006	10.46 (3.57)	2.25 (3.83)	9.84 (9.17)	-672.80 (342.14)	0.00 (83.53)
2007	9.79 (3.76)	2.06 (4.04)	10.42 (8.41)	-418.56 (323.44)	-0.02 (82.73)
2008	9.79 (3.43)	2.34 (4.68)	8.63 (7.93)	-603.76 (322.03)	0.02 (82.08)
2009	9.75 (3.21)	2.21 (4.03)	10.17 (9.37)	-454.67 (304.82)	-0.02 (79.74)
2010	8.16 (3.14)	1.98 (3.75)	10.2 (8.54)	-731.19 (288.44)	0.01 (76.35)
2011	9.57 (3.40)	2.68 (4.24)	8.51 (8.87)	-483.02 (285.00)	-0.02 (73.47)
2012	10.17 (3.48)	2.80 (4.90)	9.37 (8.56)	-541.43 (285.61)	-0.01 (74.27)
2013	9.32 (2.90)	1.20 (2.36)	9.08 (7.91)	-140.07 (233.85)	0.04 (72.95)
2014	9.67 (3.06)	2.08 (4.00)	10.44 (8.94)	-489.01 (337.77)	-0.03 (77.45)
2015	9.75 (3.23)	2.33 (4.07)	10.21 (9.15)	-429.01 (337.43)	-0.01 (76.10)
2016	9.09 (2.99)	1.61 (3.04)	11.52 (9.02)	-225.58 (336.79)	-0.01 (81.04)
2017	9.59 (3.19)	1.74 (2.99)	11.32 (10.03)	-476.58 (380.43)	0.01 (81.14)
2018	8.86 (2.60)	0.79 (2.03)	13.69 (8.56)	-527.02 (331.02)	-0.02 (80.96)

$$
\begin{aligned}
CH_{K,i} &= \sum_{k=1}^{K} HDD_{d(t_i)-k}, \quad HDD_d = \max(TT - \bar{T}_d, 0) \\
CC_{K,i} &= \sum_{k=1}^{K} CDD_{d(t_i)-k}, \quad CDD_d = \max(\bar{T}_d - TT, 0)
\end{aligned}
\tag{31}
$$

where $d(t_i)$ is the day of the ith price event, TT is a threshold temperature, in this case set to 18 degrees Celsius and $\bar{T}_d$ is the average temperature on the day d. $K = 3$ is used.

The inter-connector flow between the region of New South Wales and Queensland (QNI) is another important variable when it comes to price effects. If the QNI is positive, the flow is north into Queensland from New South Wales. A positive flow indicates that the Queensland system is under some stress, which could lead to a longer lasting episodes with increased prices. Also an unexpected load is likely to be associated with increased prices. The difference between the actual load and an estimate of the load from an autoregressive model is used as a measure of unexpected demand (UL). Since the time between any kind of price events is examined and duration can be assumed to depend on the length of the preceding event, this variable is also included as a covariate.

Summary statistics for the covariates are presented in Table 3. There do not seem to be any extreme variations in temperature or unexpected load over the years, the QNI on the other hand shows a somewhat more volatile behaviour.

3.1 Tests and estimation results

The results from the estimation of an ordinary Weibull survival model, that is, examining duration model when y_i is linear, is presented in the first column of Table

4. The estimated baseline hazard is less than one, indicating a decreased hazard rate with duration time. The effect of both the cumulative number of heating and cooling degree days are positive, as expected. The daily temperature range is found to have a weak negative effect on the rate at which price events occur. Since the difference between the daily maximum and minimum temperatures is smaller in the summer, this could reflect the fact that the system operates much closer to capacity than during winter. The effect of unexpected load is positive. The higher the unexpected load, the higher is the rate at which price events occur. The same is true when is comes to QNI, as a positive flow indicates some stress to the Queensland system. Finally, the length of the last event seems to have a positive effect on the rate of price event occurrences.

The null hypothesis of linearity is tested against an alternative of a full non-linear model with one transition. With a resulting LM-type statistic of $322 > \chi^2_{7,0.05} = 14$, the null is rejected. The full non-linear smooth transition duration model is then estimated. First, a grid search determines the starting values to avoid to get stuck on local minima. This is done by a search over a set of fixed parameters (η, c) and the set maximizing the log-likelihood is chosen as starting values. The result of the full non-linear model with one transition (i) is presented in the second column of Table 4.

All parameter estimates of the covariates in the linear part have the same sign and are rather unchanged, except from the temperature range now being non-significant. The intercept estimate is positive indicating that the rate at which price events occur seems to be increasing with the transition, all covariates constant. The result of the contribution of cooling degree days, QNI and the last event to the hazard function indicates also a significant change over the years. We perform a LR test and the null hypothesis that only the intercept is changing with the transition against the alternative of a full non-linear model, is rejected. The estimated mid-point of the transition is at $c = 0.86$, that is in January 2016. The estimate of the slope parameter is $\eta = 5.38$, which corresponds to a smooth transition over about a year, from August 2015 until June 2016. There seems in other words to occur an increase in the rate at which price events occur in the beginning of 2016.

A formal test of the null of a single transition against the alternative of two transitions is conducted. With a test statistic of $169 > \chi^2_{7,0.05} = 14$, the null of only one transition is rejected. An intercept only (ii) and a full two-transition model (iii) are therefore estimated and the results are presented in the third and fourth column of Table 4.

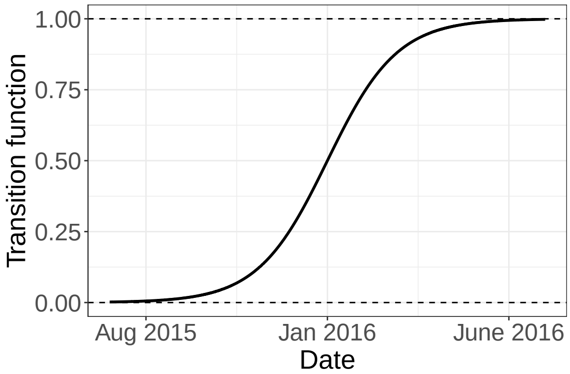
With $c_1 = 0.34$, the second transition seems to take place between August 2006 and June 2007, that is, a few months before the deregulation. The hypothesis of that the impact of all the covariates change with the second transition cannot be rejected, but with only the last event variable being significant. Comparing log-likelihoods, AIC and BIC for the models with two transitions, it seems like model (ii) is the most relevant in capturing the changing market conditions, with an increase in the rate at which price events occur in the beginning of 2007.

Continuing with a test of a third transition, the null can again not be rejected. The result of a three transition model (iv) is presented in the fifth column of Table 4. The estimation results indicate a third transition with a midpoint in August 2007 where the rate at which price events occur seems to decrease as compared to the second regime. This could indicate a move to market conditions before the deregulations. The three transition functions from this model are presented in Figure 2a, 2b and 2c.

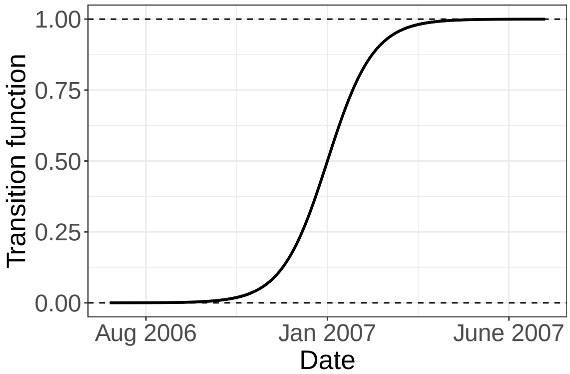
The survival function for model (iv) is presented in Figure 3, where the different

Table 4: Parameter estimates

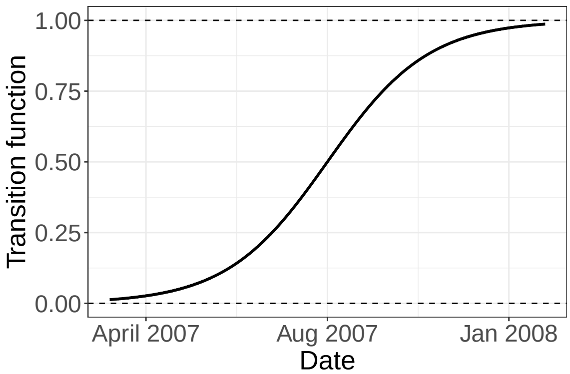
	Weibull model	STD Weibull model			
		(i)	(ii)	(iii)	(iv)
α	0.56 (0.0074)	0.61 (0.0084)	0.63 (0.0089)	0.63 (0.0097)	0.64 (0.0088)
Constant, β_0	-2.32 (0.0441)	-3.04 (0.0616)	-3.4 (0.0787)	-3.37 (0.086)	-3.51 (0.0786)
CH	0.11 (0.0226)	0.18 (0.0291)	0.18 (0.0288)	0.22 (0.0491)	0.21 (0.0286)
CC	0.1 (0.024)	0.22 (0.0332)	0.24 (0.0334)	0.18 (0.0595)	0.29 (0.0333)
TR	-0.04 (0.022)	-0.01 (0.029)	0.04 (0.0294)	0 (0.0493)	0.02 (0.0296)
UL	0.18 (0.0172)	0.19 (0.0232)	0.19 (0.0231)	0.15 (0.0399)	0.19 (0.0233)
QNI	0.34 (0.0195)	0.52 (0.0273)	0.53 (0.0282)	0.56 (0.0488)	0.59 (0.0291)
LE	0.06 (0.0131)	0.08 (0.0157)	0.09 (0.0153)	0.03 (0.0324)	0.05 (0.0166)
Constant, $\psi_{1,0}$		1.17 (0.057)	0.97 (0.0745)	0.99 (0.1036)	1.18 (0.0706)
CH_1		0.06 (0.0503)	0.06 (0.0515)	0.09 (0.056)	0.04 (0.0508)
CC_1		-0.16 (0.0487)	-0.18 (0.0498)	-0.21 (0.0553)	-0.22 (0.0493)
TR_1		-0.05 (0.0452)	-0.1 (0.0467)	-0.12 (0.0529)	-0.08 (0.0463)
UL_1		-0.04 (0.0352)	-0.04 (0.0359)	-0.06 (0.04)	-0.04 (0.0358)
QNI_1		-0.31 (0.0405)	-0.32 (0.0421)	-0.31 (0.0485)	-0.37 (0.0421)
LE_1		0.07 (0.0283)	0.06 (0.0287)	0.03 (0.0297)	0.11 (0.029)
η_1		5.38 (0.3892)	5.39 (0.4745)	5.38 (0.623)	5.38 (0.4246)
c_1		0.86 (0.0037)	0.86 (0.0058)	0.86 (0.0092)	0.86 (0.0046)
Constant, $\psi_{2,0}$			0.45 (0.0574)	0.39 (0.0863)	1.21 (0.127)
CH_2				-0.06 (0.0604)	
CC_2				0.11 (0.0719)	
TR_2				0.05 (0.0618)	
UL_2				0.06 (0.0493)	
QNI_2				-0.03 (0.0609)	
LE_2				0.09 (0.0368)	
η_1			5.8 (0.6034)	5.9 (0.5796)	5.8 (0.8379)
c_1			0.34 (0.0052)	0.34 (0.004)	0.35 (0.002)
Constant $\psi_{3,0}$					-0.87 (0.1263)
η_2					5.01 (0.3312)
c_2					0.38 (0.004)
Log-likelihood	-14 554	-14 292	-14 252	-14 243	-14 184
AIC	29 124	28 617	28 545	28 538	28 414
BIC	29 208	28 798	28 757	28 814	28 658



(a) The first transition



(b) The second transition



(c) The third transition

Figure 2: The three estimated transitions in the smooth transition model (iv)

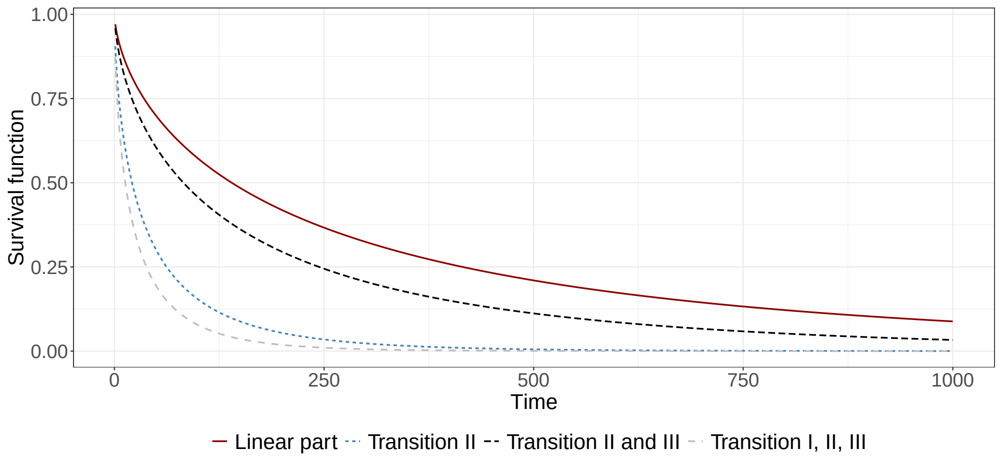


Figure 3: Survival functions from model (iv) using the mean of the covariates.

regimes are compared. It is clear that the probability of a duration to hold until a certain time point decreases with the transition to the second regime in late 2006, with a succeeding increase in 2007. After 2015, when all transitions in model (iv) have occurred, the probability decreases again.

To examine the estimated model further, we compare it with models with a deterministic structural break included sequentially for each month, instead of a smooth transition, see Figure 4a, 4b and 4c. The left axis and the solid line represents the log-likelihood of the deterministic structural break models and the dotted line (right axis) is the specific smooth transition function found by estimating model (iv) in Table 4. Figure 4a represents the log-likelihood from estimating a linear model with a dummy variable being zero up until the specified month and one thereafter, for each of the months between January 2002 to July 2017. Figure 4b and 4c corresponds to the log-likelihood from models including a deterministic structural break in January 2016 and in both January 2016 and July 2007 respectively, and with an additional break included as described above.

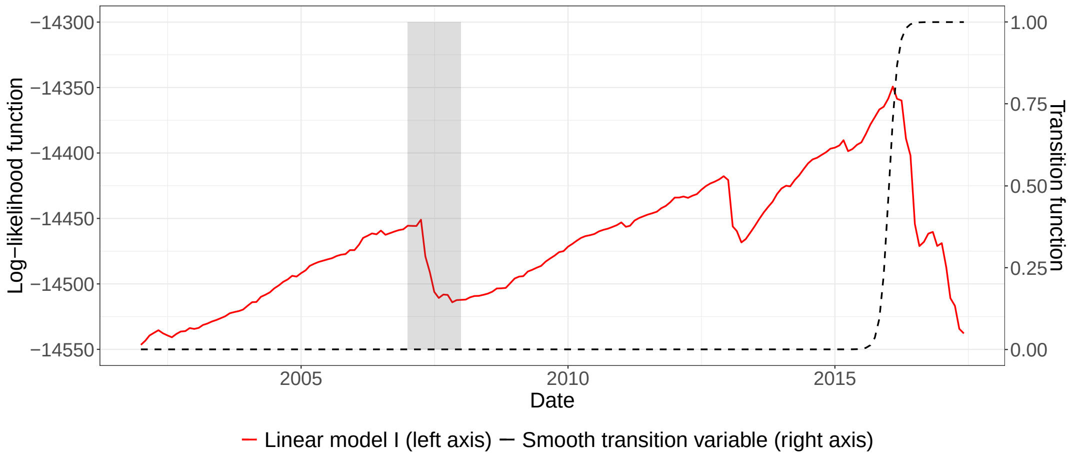
It is clear that the smooth transition model to some extent captures the three breaks maximizing the log-likelihood of the linear model. The second transition function seem to indicate a transition somewhat before the linear model with a maximized log-likelihood given a structural break in April 2007. The estimation results from the three linear models with deterministic structural breaks maximizing the log-likelihoods are presented in 6. The estimated parameters show an increase in the hazard at the first two time points and a decrease in the third, with about the same magnitudes as in the smooth transition models.

In all of the above specifications, power dependence of the hazard is assumed and estimated to be less that one. This implies a decreased hazard rate with duration time. As a check of the assumption of power dependence, a comparison with a constant baseline hazard is in addition examined but there seems to be adequate to assume an, in this case, decreasing power dependence.

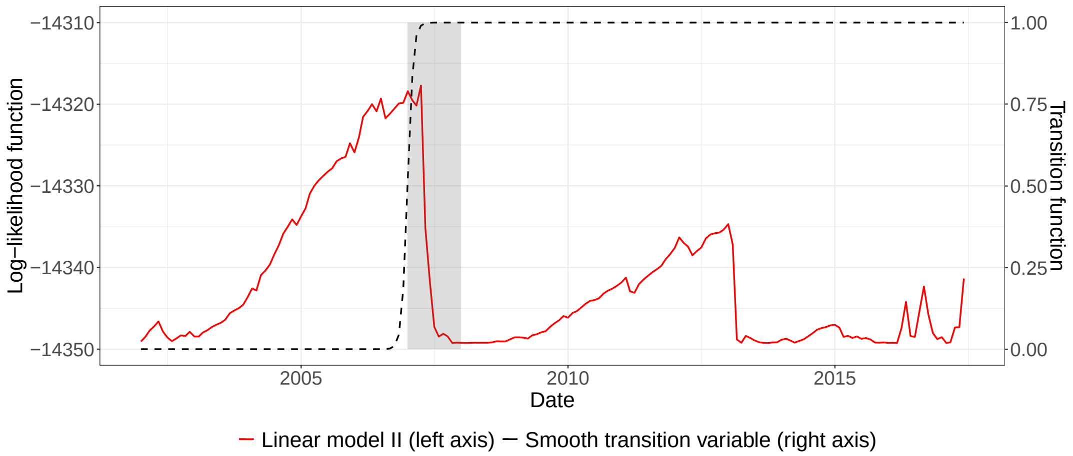
4 Conclusions

In this paper, we introduce a smooth transition duration model, allowing for both time-varying covariates and for a duration time to vary with smooth transitions over different regimes. The proposed LM-tests are shown to have the desired properties

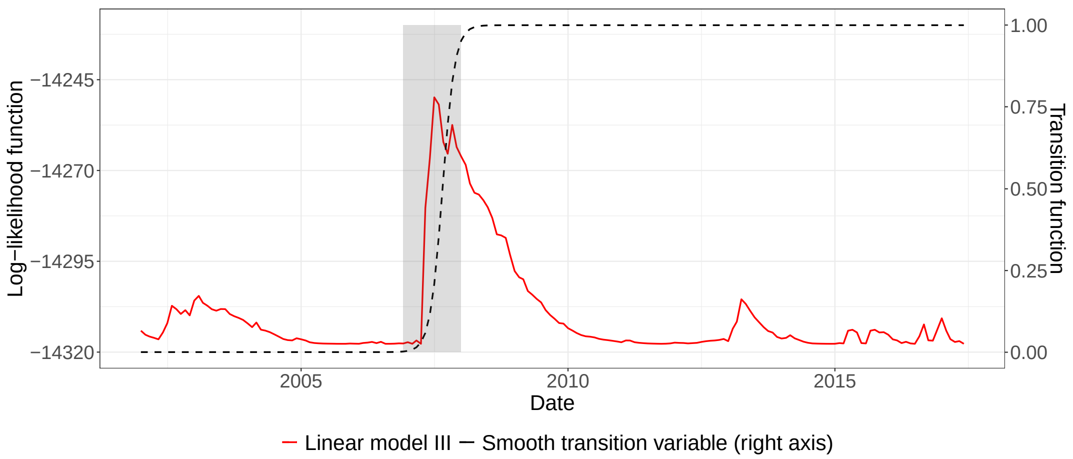
In our empirical application, we model time until an unexplained, short price



(a) The first transition and linear models with one structural break



(b) The second transition and linear models with two structural breaks



(c) The third transition and linear models with three structural breaks

Figure 4: The three estimated transitions in the smooth transition model (iv). The shaded area correspond to 2007.

increase, i.e. an abnormal price event, in the electricity spot price in Queensland, Australia. By modeling this duration time, we investigate if there is a change over time and if, more specifically, the behaviour by market participants is affected by regulatory changes. The results show that there seems to be more than one transition over the period of study. The 'first' transition is found to take place in the beginning of 2016, while the 'second' is found in the first half of 2007, both indicating an increase in the rate at which price events occur. As the deregulation in the Australian electricity market began in 2007, with energy prices in South East Queensland deregulated in 2016, the findings are in line with what could be expected.

The structural changes seem, in addition, not to be instantaneous but rather smooth over a period of time, justifying the use of the proposed model. The results are in line with the conclusions of (Hurn et al., 2016), who find support for the statement that the deregulation altered the behaviour of market participants.

Appendix

Table 5: Size adjusted power of the test, 10 000 replicates

Slope, γ	*Scale*	K	150	300	600	1 200	2 400	4 800
10	0.5	1	0.088	0.127	0.213	0.386	0.639	0.909
10	0.5	3	0.074	0.114	0.166	0.300	0.563	0.882
20	0.5	1	0.096	0.145	0.250	0.451	0.724	0.953
20	0.5	3	0.079	0.129	0.197	0.365	0.663	0.939
40	0.5	1	0.099	0.151	0.260	0.468	0.746	0.962
40	0.5	3	0.081	0.134	0.204	0.382	0.688	0.951

Note: The size adjusted power is presented for different sample sizes, T. The parameters under the alternative is $\boldsymbol{\psi} = (0.25, 0.25, 0.25)$ scaled with 0.5 and K is the number of parameters tested; only intercept or both intercept and impact of the two control variables.

Table 6: Model comparisons, three different deterministic structural break models

	(v)	(vi)	(vii)
α	0.61 (0.0082)	0.61 (0.0083)	0.63 (0.0085)
Constant	-2.86 (0.0539)	-3.16 (0.0676)	-3.25 (0.0686)
CH	0.17 (0.0230)	0.17 (0.0228)	0.20 (0.0229)
CC	0.12 (0.0236)	0.13 (0.0236)	0.19 (0.0242)
TR	-0.03 (0.0222)	0.00 (0.0224)	0.00 (0.0224)
UL	0.17 (0.0172)	0.17 (0.0172)	0.16 (0.0173)
QNI	0.36 (0.0195)	0.37 (0.0196)	0.37 (0.0196)
LE	0.10 (0.0129)	0.11 (0.0125)	0.08 (0.0124)
D_{Jan16}	0.83 (0.0403)	0.70 (0.0428)	0.82 (0.0448)
D_{Apr07}		0.41 (0.0523)	1.43 (0.0919)
D_{Jul07}			-1.14 (0.0878)
Log-likelihood	$-14\,349$	$-14\,318$	$-14\,250$
AIC	28 717	28 655	28 522
BIC	28 812	28 762	28 638

With the assumptions above fulfilled, the following two theorems can be proved:

Theorem 1. Assume that Assumption 1-5 holds. The maximum likelihood estimator $(\hat{\alpha}, \hat{\boldsymbol{\theta}})$ is consistent for $(\alpha_0, \boldsymbol{\theta}_0)$. That is,

$$(\hat{\alpha}, \hat{\boldsymbol{\theta}}) \xrightarrow{\text{P}} (\alpha_0, \boldsymbol{\theta}_0) \tag{32}$$

as $N \to \infty$.

The consistency of the estimators is proved by verifying the conditions of Theorem 2.5 in Newey and McFadden (1994)

i. Identification. If $\theta \neq \theta_0$, then $l(\theta) \neq l(\theta_0)$

ii. The parameter space is a compact space

iii. $l(\theta)$ is continuous at each $\theta \in \Theta$ with probability 1

iv. E $| \, l(\theta) \, | < \infty$.

The identification in (i.) can be proved as follows. First note that Hurn et al. (2016) prove that $\boldsymbol{\beta}'\mathbf{x} \neq \boldsymbol{\beta}_0'\mathbf{x}$ for $\boldsymbol{\beta} \neq \boldsymbol{\beta}_0$, $\boldsymbol{\psi}'\mathbf{x} \neq \boldsymbol{\psi}_0'\mathbf{x}$ for $\boldsymbol{\psi} \neq \boldsymbol{\psi}_0$, $\gamma\boldsymbol{\psi}'\mathbf{x} \neq \gamma_0\boldsymbol{\psi}'\mathbf{x}$ for $\gamma \neq \gamma_0$ and $\mathbf{c}\boldsymbol{\psi}'\mathbf{x} \neq \mathbf{c}_0\boldsymbol{\psi}'\mathbf{x}$ for $\mathbf{c} \neq \mathbf{c}_0$. From this and the strict monotonicity it follows that $(\boldsymbol{\beta} + \boldsymbol{\psi}G(s_i; \mathbf{c}, \gamma))'\,\mathbf{x} \neq (\boldsymbol{\beta}_0 + \boldsymbol{\psi}_0 G(s_i; \mathbf{c}_0, \gamma_0))'\,\mathbf{x}$ when $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ and consequently $\exp\left((\boldsymbol{\beta} + \boldsymbol{\psi}G(s_i; \mathbf{c}, \gamma))'\,\mathbf{x}\right) \neq \exp\left((\boldsymbol{\beta}_0 + \boldsymbol{\psi}_0 G(s_i; \mathbf{c}_0, \gamma_0))'\,\mathbf{x}\right)$. Hence, it follows that, as argued by Hurn et al. (2016) and Newey and McFadden (1994), that for the contribution to the likelihood from each observation $l_i(\boldsymbol{\theta}) \neq l_i(\boldsymbol{\theta})$ for $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ and the model is identified. Point (ii.) holds by Assumption A1. Point (iii.) is trivially true as both $(\boldsymbol{\beta} + \boldsymbol{\psi}G(s_i; \mathbf{c}, \gamma))'\,\mathbf{x}$ and exp are continuous functions. The last point (iv.) follows from

$$\begin{aligned} E|l_i(\alpha, \boldsymbol{\theta})| &\leq E|c_i \ln(\alpha)| + E|c_i(\alpha-1)\ln(t_i)| + E|c_i y_i| + E|t_i^{\alpha} e^{y_i}| + E|r_i^{\alpha} e^{y_i}| && (33) \\ &\leq |\ln(\alpha)| + E|(\alpha-1)\ln(t_i)| + |y_i| + E|t_i^{\alpha} e^{y_i}| + |r_i^{\alpha} e^{y_i}| && (34) \\ &\leq |\ln(\alpha)| + |(\alpha-1)|E|\ln(t_i)| + |y_i| + e^{y_i} E|t_i^{\alpha}| + e^{y_i} r_i^{\alpha} && (35) \end{aligned}$$

which is bounded as $E|\ln(t_i)| < \infty$ and $E|t_i^{\alpha}| < \infty$ (see e.g. Lehman (1963)).

Theorem 2. Assume that Assumption 1-5 holds and that $(\hat{\alpha}, \hat{\boldsymbol{\theta}}) \xrightarrow{\text{P}} (\alpha_0, \boldsymbol{\theta}_0)$ as $N \to \infty$, then

$$\sqrt{T}((\hat{\alpha}, \hat{\boldsymbol{\theta}}) - (\alpha_0, \boldsymbol{\theta}_0)) \xrightarrow{\text{D}} N(0, \mathbf{J}^{-1}) \tag{36}$$

where $\mathbf{J} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \mathbf{J}_i$. To prove asymptotic normality of the estimators, the following is needed in addition to the conditions for consistency

i. θ_0 is an interior of Θ

ii. $l(\theta)$ is twice continuously differentiable and $l(\theta) > 0$ in a neighbourhood $\mathcal{N}$ of θ_0

iii. $\int \sup_{\theta \in \mathcal{N}} || \frac{\partial l(\theta)}{\partial(\theta)} || \partial u < \infty$ where $u = (y, x\prime)\prime$

iv. $J = \lim_{T\to\infty} (1/T) \sum_{i=1}^{T} E\left[\frac{\partial l(\theta)}{\partial \theta} \frac{\partial l(\theta)}{\partial \theta\prime}\right]$ exists and is non-singular

v. $E\text{sup}_{\theta\in\mathcal{N}}||\frac{\partial^2 l(\theta)}{\partial\theta\theta'}|| < \infty$

The condition in (i.) follows from Assumption 1. Point (ii.) holds due to properties of the model. Point (iii.) is proven by (Hurn et al., 2016, P. 729).

Point (iv.) is proved as follows. From assumption A3. we have that $E\mathbf{x}\mathbf{x}'$ is positive definite. Let $f(\mathbf{x})$ denote the pdf of $\mathbf{x}$ and $g(\mathbf{x}) = (c_i - t_i^\alpha e^{y_i} + r_i^\alpha e^{y_i})^2$ where $g(\mathbf{x}) \geq 0$. The expectation is

$$E(g(\mathbf{x})\mathbf{x}\mathbf{x}') \quad = \quad \int g(\mathbf{x})\mathbf{x}\mathbf{x}'p(\mathbf{x})dx$$

where $p(\mathbf{x})$ is the pdf of $\mathbf{x}$. Now $q(\mathbf{x}) = g(\mathbf{x})p(\mathbf{x})/K$ where K is a normalizing constant (which exists due to Assumption A3.), is a proper pdf and hence

$$E(g(\mathbf{x})\mathbf{x}\mathbf{x}') \quad = \quad K\int \mathbf{x}\mathbf{x}'q(\mathbf{x})dx$$

which imply that $Eg(\mathbf{x})\mathbf{x}\mathbf{x}'$ is positive definite. Next, as above, let $a_i = c_i(\alpha^{-1} + \ln(t_i)) - t_i^\alpha e^{y_i}\ln(t_i) + r_i^\alpha e^{y_i}\ln(r_i)$

$$\begin{aligned} J_{N_n}(\alpha,\boldsymbol{\theta}) &= \frac{1}{N_n}\sum_{i=1}^{N_n} E\left[\begin{array}{cc} a_i^2 & a_i\left(c_i - t_i^\alpha e^{y_i} + r_i^\alpha e^{y_i}\right)\mathbf{x}'\mathbf{K}'(r) \\ & \left(c_i - t_i^\alpha e^{y_i} + r_i^\alpha e^{y_i}\right)^2\mathbf{K}(r)\mathbf{x}\mathbf{x}'\mathbf{K}'(r))' \end{array}\right] \\ &= \end{aligned}$$

As $a_i^2 > 0$ and the result above we have positive definiteness.

$$J_N(\alpha,\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N} E\left[\begin{array}{cc} a_i^2 & a_i\left(c_i - t_i^\alpha e^{y_i} + r_i^\alpha e^{y_i}\right)\mathbf{w}'(s_i) \\ & \left(c_i - t_i^\alpha e^{y_i} + r_i^\alpha e^{y_i}\right)^2\mathbf{w}(s_i)\mathbf{w}(s_i)' \end{array}\right] \tag{37}$$

Point (v.) is fulfilled by noting that the individual elements are bounded. The derivatives are given below.

$$\begin{aligned}
\frac{\partial l\left(\alpha,\boldsymbol{\theta}\right)}{\partial\alpha}\frac{\partial l\left(\alpha,\boldsymbol{\theta}\right)}{\partial\alpha} &= \sum_i \left(c_i(\hat{\alpha}^{-1} + \ln(t_i)) - t_i^{\hat{\alpha}}e^{\hat{y}_i}\ln(t_i) + r_i^{\hat{\alpha}}e^{\hat{y}_i}\ln(r_i)\right)^2 \\
\frac{\partial l\left(\alpha,\boldsymbol{\theta}\right)}{\partial\boldsymbol{\phi}_1}\frac{\partial l\left(\alpha,\boldsymbol{\theta}\right)}{\partial\boldsymbol{\phi}_1'} &= \sum_i \left(c_i - t_i^{\hat{\alpha}}e^{\hat{y}_i^A} + r_i^{\hat{\alpha}}e^{\hat{y}_i^A}\right)^2\mathbf{x}_i\mathbf{x}_i' \\
\frac{\partial l\left(\alpha,\boldsymbol{\theta}\right)}{\partial\boldsymbol{\phi}_2}\frac{\partial l\left(\alpha,\boldsymbol{\theta}\right)}{\partial\boldsymbol{\phi}_2'} &= \sum_i \left(c_i - t_i^{\hat{\alpha}}e^{\hat{y}_i^A} + r_i^{\hat{\alpha}}e^{\hat{y}_i^A}\right)^2\mathbf{x}_i\mathbf{x}_i's_i^2 \\
\frac{\partial l\left(\alpha,\boldsymbol{\theta}\right)}{\partial\alpha}\frac{\partial l\left(\alpha,\boldsymbol{\theta}\right)}{\partial\boldsymbol{\phi}_1'} &= \sum_i \left(c_i(\hat{\alpha}^{-1} + \ln(t_i)) - t_i^{\hat{\alpha}}e^{\hat{y}_i}\ln(t_i) + r_i^{\hat{\alpha}}e^{\hat{y}_i}\ln(r_i)\right)\left(c_i - t_i^{\hat{\alpha}}e^{\hat{y}_i^A} + r_i^{\hat{\alpha}}e^{\hat{y}_i^A}\right)\mathbf{x}_i' \\
\frac{\partial l\left(\alpha,\boldsymbol{\theta}\right)}{\partial\alpha}\frac{\partial l\left(\alpha,\boldsymbol{\theta}\right)}{\boldsymbol{\phi}_2'} &= \sum_i \left(c_i(\hat{\alpha}^{-1} + \ln(t_i)) - t_i^{\hat{\alpha}}e^{\hat{y}_i}\ln(t_i) + r_i^{\hat{\alpha}}e^{\hat{y}_i}\ln(r_i)\right)\left(c_i - t_i^{\hat{\alpha}}e^{\hat{y}_i^A} + r_i^{\hat{\alpha}}e^{\hat{y}_i^A}\right)\mathbf{x}_i's_i \\
\frac{\partial l\left(\alpha,\boldsymbol{\theta}\right)}{\partial\boldsymbol{\phi}_1'}\frac{\partial l\left(\alpha,\boldsymbol{\theta}\right)}{\partial\boldsymbol{\phi}_2'} &= \sum_i \left(c_i - t_i^{\hat{\alpha}}e^{\hat{y}_i^A} + r_i^{\hat{\alpha}}e^{\hat{y}_i^A}\right)^2\mathbf{x}_i\mathbf{x}_i's_i
\end{aligned} \tag{38}$$

We have

$$
\begin{aligned}
\frac{\partial^2 l\left(\alpha,\theta\right)}{\partial\alpha\,\partial\alpha} &= \sum_i c_i\hat{\alpha}^{-2} - t_i^{\hat{\alpha}} e^{\hat{y}_i^A} \ln^2 t_i + r_i^{\hat{\alpha}} e^{\hat{y}_i^A} \ln^2 r_i \\
\frac{\partial^2 l\left(\alpha,\theta\right)}{\partial\theta_1\,\partial\theta_1} &= \sum_i - t_i^{\hat{\alpha}} e^{\hat{y}_i^A} x_i' x_i' + r_i^{\hat{\alpha}} e^{\hat{y}_i^A} x_i' x_i' \\
\frac{\partial^2 l\left(\alpha,\theta\right)}{\partial\theta_2\,\partial\theta_2'} &= \sum_i - t_i^{\hat{\alpha}} e^{\hat{y}_i^A} x_i' s_i + r_i^{\hat{\alpha}} e^{\hat{y}_i^A} x_i' s_i \\
\frac{\partial^2 l\left(\alpha,\theta\right)}{\partial\alpha\,\partial\theta_1'} &= \sum_i - t_i^{\hat{\alpha}} e^{\hat{y}_i^A} \ln t_i x_i' + r_i^{\hat{\alpha}} e^{\hat{y}_i^A} \ln r_i x_i' \\
\frac{\partial^2 l\left(\alpha,\theta\right)}{\partial\alpha\,\theta_2'} &= \sum_i - t_i^{\hat{\alpha}} e^{\hat{y}_i^A} \ln t_i x_i' s_i + r_i^{\hat{\alpha}} e^{\hat{y}_i^A} \ln r_i x_i' s_i \\
\frac{\partial^2 l\left(\alpha,\theta\right)}{\partial\theta_1\theta_2'} &= \sum_i - t_i^{\hat{\alpha}} e^{\hat{y}_i^A} x_i' x_i' s_i + r_i^{\hat{\alpha}} e^{\hat{y}_i^A} x_i' x_i' s_i
\end{aligned} \tag{39}
$$

$$
\mathbf{D}_{i|H_0} = \begin{bmatrix} \mathbf{D}_{11i} & \mathbf{D}_{12i} \\ \mathbf{D}_{21i} & \mathbf{D}_{22i} \end{bmatrix}_{|H_0} \tag{40}
$$

$$
\begin{aligned}
\mathbf{D}_{i|H_0} &= \begin{bmatrix} \mathbf{D}_{11i} & \mathbf{D}_{12i} \\ \mathbf{D}_{21i} & \mathbf{D}_{22i} \end{bmatrix}_{|H_0} \\
&= \left[\begin{array}{cc|c}
\frac{\partial l_i(\alpha,\phi)}{\partial\alpha}\frac{\partial l_i(\alpha,\phi)}{\partial\alpha} & \frac{\partial l_i(\alpha,\phi)}{\partial\phi_1}\frac{\partial l_i(\alpha,\phi)}{\partial\alpha} & \frac{\partial l_i(\alpha,\phi)}{\partial\phi_2}\frac{\partial l_i(\alpha,\phi)}{\partial\alpha} \\
\frac{\partial l_i(\alpha,\phi)}{\partial\alpha}\frac{\partial l_i(\alpha,\phi)}{\partial\phi_1} & \frac{\partial l_i(\alpha,\phi)}{\partial\phi_1}\frac{\partial l_i(\alpha,\phi)}{\partial\phi_1'} & \frac{\partial l_i(\alpha,\phi)}{\partial\phi_2}\frac{\partial l_i(\alpha,\phi)}{\partial\phi_1'} \\
\hline
\frac{\partial l_i(\alpha,\phi)}{\partial\alpha}\frac{\partial l_i(\alpha,\phi)}{\partial\phi_2} & \frac{\partial l_i(\alpha,\phi)}{\partial\phi_1'}\frac{\partial l_i(\alpha,\phi)}{\partial\phi_2} & \frac{\partial l_i(\alpha,\phi)}{\partial\phi_2}\frac{\partial l_i(\alpha,\phi)}{\partial\phi_2'}
\end{array}\right]_{|H_0} \\
&= \left[\begin{array}{c|cc}
\frac{\partial l_i(\alpha,\phi)}{\partial\alpha}\frac{\partial l_i(\alpha,\phi)}{\partial\alpha} & \frac{\partial l_i(\alpha,\phi)}{\partial\phi_1}\frac{\partial l_i(\alpha,\phi)}{\partial\alpha} & \frac{\partial l_i(\alpha,\phi)}{\partial\phi_2}\frac{\partial l_i(\alpha,\phi)}{\partial\alpha} \\
\hline
\frac{\partial l_i(\alpha,\phi)}{\partial\alpha}\frac{\partial l_i(\alpha,\phi)}{\partial\phi_1} & \frac{\partial l_i(\alpha,\phi)}{\partial\phi_1}\frac{\partial l_i(\alpha,\phi)}{\partial\phi_1'} & \frac{\partial l_i(\alpha,\phi)}{\partial\phi_2}\frac{\partial l_i(\alpha,\phi)}{\partial\phi_1'} \\
\frac{\partial l_i(\alpha,\phi)}{\partial\alpha}\frac{\partial l_i(\alpha,\phi)}{\partial\phi_2} & \frac{\partial l_i(\alpha,\phi)}{\partial\phi_1'}\frac{\partial l_i(\alpha,\phi)}{\partial\phi_2} & \frac{\partial l_i(\alpha,\phi)}{\partial\phi_2}\frac{\partial l_i(\alpha,\phi)}{\partial\phi_2'}
\end{array}\right]_{|H_0}
\end{aligned} \tag{41}
$$

References

Castro, V. (2013), 'The duration of business cycle expansions and contractions: are there change-points in duration dependence?', *Empirical Economics* **44**(2), 511–544.

Dijk, D. v., Teräsvirta, T. and Franses, P. H. (2002), 'Smooth transition autoregressive models – a survey of recent developments', *Econometric reviews* **21**(1), 1–47.

Hillebrand, E., Medeiros, M. C. and Xu, J. (2013), 'Asymptotic Theory for Regressions with Smoothly Changing Parameters', *Journal of Time Series Econometrics* **5**(2), 133–162.

Hurn, S., Silvennoinen, A. and Teräsvirta, T. (2016), 'A smooth transition logit model of the effects of deregulation in the electricity market', *Journal of Applied Econometrics* **31**(4), 707–733.

Kiefer, N. M. (1988), 'Economic duration data and hazard functions', *Journal of economic literature* **26**(2), 646–679.

Lara-Porras, A.-M., Alvarez, E., Garc, J. and Quesada-Rubio, J.-M. (2005), Weibull survivals with changepoints and heterogeneity, *in* J. Janssen and P. Lenca, eds, 'Proceedings of the conference on applied stochastic models and data analysis of the quantitative methods in business and industry society', ENST Bretagne, pp. 1218–1223.

Lehman, E. (1963), 'Shapes, moments and estimators of the weibull distribution', *IEEE Transactions on Reliability* **R-12**(3), 32–38.

Luukkonen, R., Saikkonen, P. and Teräsvirta, T. (1988), 'Testing linearity against smooth transition autoregressive models', *Biometrika* **75**(3), 491–499.

Newey, W. K. and McFadden, D. (1994), Chapter 36 large sample estimation and hypothesis testing, Vol. 4 of *Handbook of Econometrics*, Elsevier, pp. 2111 – 2245.

Pons, O. (2003), 'Estimation in a cox regression model with a change-point according to a threshold in a covariate', *Annals of Statististics* **31**(2), 442–463.

Teräsvirta, T. (1994), 'Specification, estimation, and evaluation of smooth transition autoregressive models', *Journal of the American Statistical Association* **89**(425), 208–218.

Teräsvirta, T., Tjøstheim, D. and Granger, C. W. J. (2010), *Modelling Nonlinear Economic Time Series*, Advanced Texts in Econometrics, Oxford University Press.

Van Den Berg, G. J. (2001), Chapter 55 - duration models: Specification, identification and multiple durations, Vol. 5 of *Handbook of Econometrics*, Elsevier, pp. 3381 – 3460.

Xiaolong, L. and Boyett, J. M. (1997), 'Estimations of a threshold parameter in cox regression', *Communications in Statistics - Theory and Methods* **26**(10), 2329–2346.