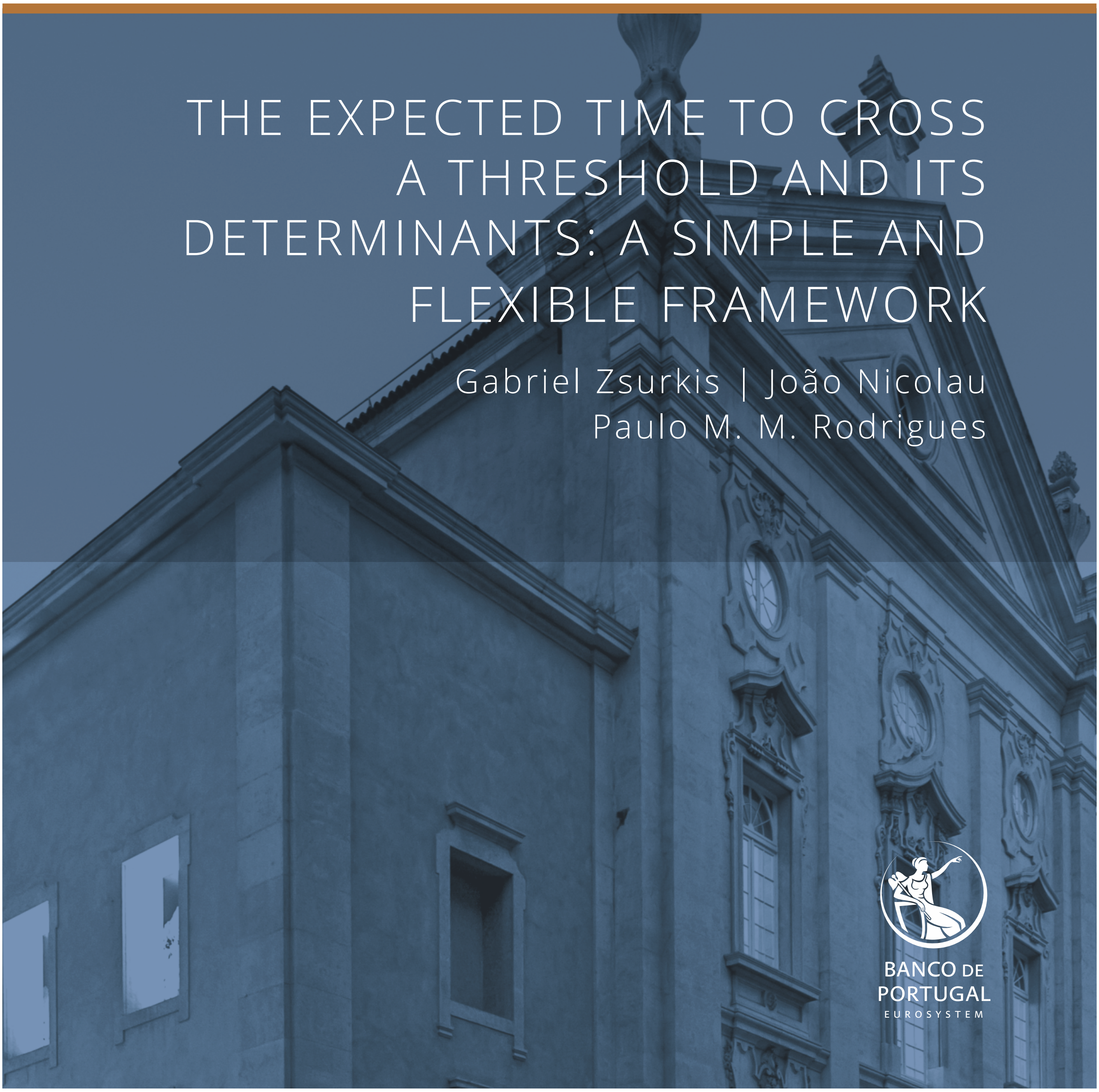
06

WORKING PAPERS 2020



THE EXPECTED TIME TO CROSS A THRESHOLD AND ITS DETERMINANTS: A SIMPLE AND FLEXIBLE FRAMEWORK

Gabriel Zsurkis | João Nicolau
Paulo M. M. Rodrigues

BANCO DE PORTUGAL
EUROSYSTEM

06

WORKING PAPERS 2020

THE EXPECTED TIME TO CROSS A THRESHOLD AND ITS DETERMINANTS: A SIMPLE AND FLEXIBLE FRAMEWORK

Gabriel Zsurkis | João Nicolau
Paulo M. M. Rodrigues

MAY 2020

Please address correspondence to
Banco de Portugal, Economics and Research Department
Av. Almirante Reis, 71, 1150-012 Lisboa, Portugal
Tel.: +351 213 130 000, email: estudos@bportugal.pt

BANCO DE PORTUGAL
EUROSYSTEM

Lisboa, 2020 • www.bportugal.pt

Working Papers | Lisbon 2020 • Banco de Portugal Av. Almirante Reis, 71 | 1150-012 Lisboa • www.bportugal.pt •
Edition Economics and Research Department • ISBN (online) 978-989-678-730-1 • ISSN (online) 2182-0422

The expected time to cross a threshold and its determinants: A simple and flexible framework

Gabriel Zsurkis
Banco de Portugal and
ISEG-Universidade de Lisboa

João Nicolau
ISEG-Universidade de Lisboa and
CEMAPRE

Paulo M. M. Rodrigues
Banco de Portugal and Nova School of
Business and Economics

May 2020

Abstract
In this paper we introduce a flexible framework to estimate the expected time (ET) an outcome variable takes to cross a threshold conditional on covariates. The proposed methodology makes use of the Markovian property and allows us to infer the impacts that covariates have on the ET an outcome variable takes to revert to a value of interest (for instance, its mean) given a specific starting point. An empirical application to the U.S. economy is provided, which investigates how the yield spread (YS) influences the ET the industrial production (IP) growth rate takes to return to its mean considering several initial values for the outcome variable. Our results suggest that the YS may have an important role in stimulating a faster return to desirable growth rates when the economy is in contraction or faces weak growth. Moreover, the YS also seems relevant in the presence of positive and high IP growth rates since a negative value of this variable may contribute for the IP growth rate to quickly return to below average values.

JEL: C32, C41, C51

Keywords: Expected time, Markov chains, nonlinearity, yield spread, industrial production.

Acknowledgements: Financial support from the Portuguese Science Foundation (FCT) through projects UID/MULTI/00491/2019, PTDC/EGE-ECO/28924/2017, and (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, LISBOA-01-0145-FEDER-022209), POR Lisboa (LISBOA-01-0145-FEDER-007722, LISBOA-01-0145-FEDER-022209) and POR Norte (LISBOA-01-0145-FEDER-022209) is gratefully acknowledged.

E-mail: gzsurkis@bportugal.pt; nicolau@iseg.utl.pt; pmrodrigues@bportugal.pt

1. Introduction

The first hitting time or first passage time, i.e., the time a variable takes to reach a certain value, is a fundamental concept in stochastic analysis and represents an important modeling tool in fields, such as finance, biology and life sciences.

Although there is a large literature in economics and finance addressing this topic (see, for instance, Durbin 1971, Lo *et al.* 2002 or Giesecke 2006), first-hitting time densities are mostly obtained for Wiener diffusion processes under the assumption of continuous-time, due to the tractability offered by the Itô calculus. However, this approach often requires strong computational efforts and closed form solutions are known only for some standard continuous-time models.

Since most economic and financial data is only available in discrete time, researchers usually opt for modeling duration time as a stochastic process instead of defining duration as the first time a stochastic process crosses a given threshold. Thus, continuous-time based first passage time densities and duration models (typically built in discrete time) have the same objective, namely, to characterize the length of time that separates different stochastic events. In fact, as illustrated by Whitmore (1986), duration models can be seen as reduced form representations of first passage time densities.

Most of the existing duration analysis literature is based on the specification of the hazard function, that is, on the conditional probability of exiting the initial state within a short interval having survived up to the starting time of that interval. Thus, the hazard function specification emphasizes the conditional probabilities[1]. Since a duration process can intuitively be associated with a dynamic sequence of conditional probabilities, the hazard-based approach is a convenient way to interpret duration data and can be sufficiently flexible to handle relevant issues such as the presence of censored observations and time-varying covariates. For instance, parametric hazard models have been used in labor economics to examine duration dependence and the determinants of unemployment exit probabilities (see, for instance, Meyer 1990, McCall 1994 and Sueyoshi 1995); another example is its application to firm survival (see, for instance, Audretsch and Mahmood 1995 and Mata and Portugal 2002). In addition, this methodology has also been employed (without covariates) to investigate the presence of duration dependence in economic cycles (see, for instance, Sichel 1991 and Ohn *et al.* 2004).

A closely related approach to model duration dependence is to treat the occurrence of a given event as a random variable which follows a point process[2]. Let $\{t_i\}_{i\in\{1,2,\dots\}}$, with $0 \leq t_i \leq t_{i+1}$, be a sequence of non-negative random variables representing the times at which the events occur. The sequence $\{t_i\}$ is called a point process. A complete description of such processes is formulated in terms

1. Note that, for any hazard function specification there is a mathematically equivalent representation in terms of a probability distribution; see Kiefer (1988).

2. For an introduction to point processes see, for instance, Cox and Isham (1980).

of the conditional intensity function which can be associated, roughly speaking, with the probability per unit of time, to observe an event in the next instant[3]. Thus, different parameterizations of this function result in different point process models. Existing models can be grouped into two classes. The first, formulated in calendar time, considers that the marginal effects of an event that has occurred in the past is independent of the intervening history; and the second class, focuses on the intervals between events and assumes that the duration between successive events depends on the number of intervening events. The autoregressive conditional dynamic (ACD) model proposed by Engle and Russell (1998) is an important model of this class[4].

In this paper, we focus on first hitting time processes. Thus, unlike point process models, we are not interested in the actual sequence $\{t_i\}$, but only in the random variable associated with the time at which the event occurs for the first time (t_1). As already mentioned, the first hitting time problems were mostly addressed in continuous time invoking Wiener processes, which involve complex mathematical concepts and can result in models which are difficult to estimate.

Nicolau (2017) introduces an intuitive and easy to implement framework for estimating the first passage time probability function in a discrete time context. The main contribution of the present work is the introduction of a novel approach to estimate transition probabilities allowing for covariates and which generalizes the framework of Nicolau (2017). To this end, we adapt the approach proposed by Islam and Chowdhury (2006) to estimate covariate-dependent Markov models of any order to the present context.

Understanding how a set of covariates influences the time a dependent variable takes to cross a fixed threshold may provide relevant insights on the potential causal relationships between economic variables. The proposed covariate-dependent expected time (ET) to cross a threshold estimator may also be a useful tool to support macroeconomic policy decisions, where there are desirable values or even formal targets for some key variables, such as, output growth, inflation, or unemployment. Thus, it is important to assess the effectiveness of the instruments (or covariates) in driving the outcome variable towards some preassigned values. In practice, the impact of the covariates may not be symmetric and may also depend on the distance between the starting point and the target value. Consider, for instance, the connection between monetary policy and real economic growth. Since both negative and above-trend growth rates are undesirable, monetary policy plays a key role in fostering a healthy level of economic growth. To this end, a tight monetary policy is adopted when the rapid economic growth causes inflationary pressures and an easy one is implemented in a recession in order to boost a rapid economic recovery. However, it has been noticed in the literature that

3. The conditional intensity function can be seen as a counterpart to the hazard function.

4. The ACD model and its extensions have become a leading tool in modeling irregularly spaced high-frequency financial data, which are characterized by the occurrence of strong clustering structures in the waiting times between consecutive events.

the responsiveness of the real economy to monetary shocks is different during recessions and expansions (see, for instance, Florio 2004 and references therein). The framework introduced in this paper allows us to investigate these possible nonlinear dynamics by estimating the ET conditional on different starting values. If, for a given starting value, changes in covariates are reflected in changes in the ET estimation, it suggests that the chosen covariates affect the movement towards a specified threshold in that specific situation. When other starting points are considered the conclusions may, however, differ.

To further illustrate the usefulness of the approach introduced in this paper an application to the industrial production (IP) - yield spread (YS) relationship is provided. Until the 2008-2010 financial crisis, short-term nominal interest rates were the primary monetary policy instrument used to achieve price stability, which is a key objective of central banks. For instance, in response to the financial crisis, central banks cut nominal interest rates in order to stimulate economic growth. However, as short-term interest rates in recent years have been close to their zero lower bound and economic growth remained low, unconventional monetary policies such as quantitative easing have been employed to reduce long-term interest rates and spur aggregate demand. In our analysis we will use the sovereign yield curve slope as a proxy for the monetary policy stance since it captures both conventional monetary policy and unconventional measures such as asset purchases and forward-guidance; see Saldías (2017). Thus, the focus of the empirical application is to investigate whether YS[5] influences the ET the IP growth rate takes to return to its stationary mean starting from a specific value.

The remainder of the paper is organized as follows. Section 2 introduces the proposed methodology to estimate the conditional ET to cross a threshold (given a specific starting point). Section 3 investigates the finite sample properties of the parameter estimates that describe the relationship between the ET and the covariates. Section 4 presents an empirical application to the U.S. economy, where we infer how YS influences the ET that the IP growth rate takes to return to its mean. Section 5 concludes and a Technical Appendix collects the detailed proofs of the results presented in the paper.

2. The proposed methodology

2.1. The process and probabilities of interest

Let $\{(y_t, \mathbf{x}_t)\}$ be a vector of discrete-time processes with state space $\Re$ characterized by the following Assumption.

5. YS is computed as the difference between the 10-years government bond yields and the 3-month T-bill rate.

Assumption A.

(A1) $y_t|\mathbf{x}_t$ *is a Markov process of order* r;
(A2) $\{(y_t, \mathbf{x}_t)\}$ *is a jointly stationary vector stochastic process.*

Let A be a measurable set of range D of the process of interest, and define the first hitting time of A as $T_A := \inf\{t > 0 : y_t \in A\}$. There is a σ-finite measure $m(y)$ such that $m(A) > 0$ implies $E(T_A|X_0 = a) < \infty$ for every $a \in D\backslash\overline{A}$, where $\overline{A}$ is the closure of set A. Assumption A2 ensures that the process $\{y_t\}$ is positive Harris recurrent, that is, if the process starts from a level a not belonging to the generic set A, it will visit A as $T \to \infty$ almost surely an infinite number of times (see Meyn *et al.* 2009, chapter 9).

Consider the first hitting time $T_{z_1} = \inf\{t > 0 : y_t \geq z_1\}$ and that the process starts at z_0, with $z_0 < z_1$. The case $z_0 > z_1$ with $T_{z_1} = \inf\{t > 0 : y_t \leq z_1\}$ is almost analogous[6]. The distribution of T_{z_1} is usually difficult to derive, especially for non-linear processes. Thus, we consider a simple semi-parametric method to estimate these quantities. First, we define the following binary variable:

$$S_t := \begin{cases} 0 & \text{if } \ y_t < z_1, \ y_{t-1} < z_1, ..., y_{t-k+1} < z_1, y_{t-k} \leq z_0, \\ 1 & \text{otherwise,} \end{cases} \tag{1}$$

where $k \geq 0$ and $S_0 = 0$ if $y_0 = z_0$ (note that z_0 is the starting value of the process). Then, the probability that y_t crosses the threshold z_1 for the first time starting from z_0 is,

$$P(T_{z_1} = t) = P(S_t = 1, S_{t-1} = 0, S_{t-2} = 0, ..., S_1 = 0|S_0 = 0),$$

which is equivalent to

$$P(T_{z_1} = t) = \big(1 - p_t\big) \prod_{i=1}^{t-1} p_i \tag{2}$$

where $p_i := P(S_i = 0|S_{i-1} = 0, S_{i-2} = 0, ..., S_0 = 0)$ (see Appendix for details) .

Proposition 1. *Considering that Assumption A1 holds and that $y_t|\mathbf{x}_t$ is a Markov process of order r, then $S_t|\mathbf{x}_t$ is also an r^{th} order two-state Markov chain.*

Since in view of the Markovian property if $t > r$ then $p_t(\mathbf{x}) = p_r(\mathbf{x})$, from Proposition 1 and expression (2) it follows that,

6. In practice, we can easily transform a $z_0 > z_1$ case into $z_0 < z_1$ by replacing z_0, z_1 and y_t by $-z_0$, $-z_1$ and $-y_t$, respectively.

$$P(T_{z_1} = t|\mathbf{x}) = \begin{cases} \left[1 - p_t(\mathbf{x})\right] \prod_{i=1}^{t-1} p_i(\mathbf{x}) & \text{for } t \le r, \\ \left\{ \left[1 - p_r(\mathbf{x})\right] \prod_{i=1}^{r-1} p_i(\mathbf{x}) \right\} p_r(\mathbf{x})^{t-r} & \text{for } t > r, \end{cases} \tag{3}$$

where $p_i(\mathbf{x}) = P(S_t = 0|S_{t-1} = 0, ..., S_{t-i} = 0|\mathbf{x})$ for $1 \le i \le r$.

2.2. Covariate-dependent transition probabilities

When Assumption A1 and Proposition 1 hold, we can treat $S_t|\mathbf{x}_t$ as a Markov chain with state space $\{0, 1\}$ and use standard Markov chain inference to estimate the covariate-dependent transition probabilities.

For instance, if r=1, the transition probability matrix is

$$P(\mathbf{x}) = \begin{bmatrix} \pi_{00}(\mathbf{x}) & \pi_{01}(\mathbf{x}) \\ \pi_{10}(\mathbf{x}) & \pi_{11}(\mathbf{x}) \end{bmatrix}$$

with

$$p_1(\mathbf{x}) := \pi_{00}(\mathbf{x}) = P(S_t = 0|S_{t-1} = 0; \mathbf{x}) = \frac{\exp^{\mathbf{x}'\boldsymbol{\beta}_1}}{1 + \exp^{\mathbf{x}'\boldsymbol{\beta}_1}} =: \Lambda(\mathbf{x}'\boldsymbol{\beta}_1)$$

and $\Lambda(\mathbf{x}'\boldsymbol{\beta}_1)$ is the cumulative density function of the logistic distribution.

The generalization to higher order Markov chains is straightforwardly achieved by extending the first order Markov chain model. To this end, note that the transition probabilities of the rth order model can be arranged in a $2^r \times 2$ matrix of which we only need a line; see Islam and Chowdhury (2006). As an illustration, consider a matrix with the outcomes of an rth order Markov chain S_t, i.e.,

$$\begin{array}{cccccc} m & \{S_{t-r} & S_{t-(r-1)} & \dots & S_{t-1}\} \\ 1 & 0 & 0 & \dots & 0 \\ 2 & 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 2^r - 1 & 1 & 1 & \dots & 0 \\ 2^r & 1 & 1 & \dots & 1 \end{array} \begin{bmatrix} & S_t & \\ 0 & & 1 \\ 0 & & 1 \\ \vdots & & \vdots \\ 0 & & 1 \\ 0 & & 1 \end{bmatrix},$$

where m is an index that identifies each of the possible outcomes of $\{S_{t-1}, S_{t-2}, ..., S_{t-r}\}$. For instance, $m = 1$ corresponds to the outcomes $S_{t-1} = 0, S_{t-2} = 0, ..., S_{t-r} = 0$.

As shown in (3), the covariate-dependent transition probabilities $p_1(\mathbf{x})$, ..., $p_r(\mathbf{x})$ are needed to obtain the probability function for $T_{z1}|\mathbf{x}$. Hence, if $S_t|\mathbf{x}$ is an rth order Markov chain, we can define,

$$p_r(\mathbf{x}) = P(S_t = 0|S_{t-1} = 0, ..., S_{t-r} = 0; \mathbf{x}) = \Lambda(\mathbf{x}'\boldsymbol{\beta}_r). \tag{4}$$

Moreover, for $j = 1, ..., r-1$,

$$p_j(\mathbf{x}) = P(S_t = 0|S_{t-1} = 0..., S_{t-j} = 0\,; \mathbf{x}) = \Lambda(\mathbf{x}'\boldsymbol{\beta}_j). \tag{5}$$

Then, for an rth order Markov chain, the log-likelihood function of a Markov chain of order $j < r$ will be used to estimate the parameter vector $\boldsymbol{\beta}_j$ and, consequently, obtain the probabilities $p_1(\mathbf{x})$, ..., $p_{r-1}(\mathbf{x})$.

Despite similarities between (4) and (5) and the standard logit model for binary responses, the proposed approach is less restrictive. Firstly, the focus here is on the transition probabilities between states and not on the conditional probability of success. Moreover, no specific functional form for the underlying latent variable model is assumed. In fact, the only assumption we make on the data generating process of y_t is assumption (A1).

2.3. Parameter estimation

For an ith order Markov chain the log-likelihood function can be expressed as the sum of 2^i components, where each represents a particular outcome of $\{S_{t-1}, S_{t-2}, ..., S_{t-i}\}$; see the Appendix for details. Thus, we can maximize individually the part of the log-likelihood function which corresponds to S_{t-1}=S_{t-2}=...=$S_{t-i} = 0$, considering for observation t that,

$$\begin{aligned}\ln L_i &= \ln f(S_t|S_{t-1} = 0, ..., = S_{t-i} = 0; \mathbf{x}_t; \boldsymbol{\beta}_i) \\ &= \delta_i\, S_t \ln\Big(\Lambda(1 - \mathbf{x}_t'\boldsymbol{\beta}_i)\Big) + \delta_i(1 - S_t)\ln\Big(\Lambda(\mathbf{x}_t'\boldsymbol{\beta}_i)\Big),\end{aligned} \tag{6}$$

where $f(.)$ is a conditional density function, and δ_i is an indicator function which is equal to one when $S_{t-1} = 0, ...S_{t-i} = 0$ and zero otherwise. Note that when $\delta_i = 1$ (6) corresponds to the conditional log-likelihood function of the well-known logit model (see, for instance, Hayashi 2000).

2.3.1. Consistency and asymptotic normality of the parameter estimators. Since estimation of the transition probabilities and consequently of the ET to cross a threshold only depends on $\boldsymbol{\beta}_i$, $1 \leq i \leq r$, it is crucial that consistent estimates of these coefficient vectors are obtained.

Theorem 1. ***Consistency of conditional MLE without compactness***
Let $\{S_t, \mathbf{x}_t\}$ be jointly stationary with conditional density $f(S_t|S_{t-1} = ... = S_{t-i} = 0; \mathbf{x}_t; \boldsymbol{\beta}_i)$ and

$$\hat{\boldsymbol{\beta}}_i = \underset{\boldsymbol{\beta}_i \in B_i}{argmax}\ \frac{1}{T}\sum_{t=1}^{T} \ln f(S_t|S_{t-1} = ... = S_{t-i} = 0; \mathbf{x}_t; \boldsymbol{\beta}_i) \tag{7}$$

the quasi-ML estimator. Moreover, consider that,

(1) *the true parameter vector $\boldsymbol{\beta}_i$ is an element of the interior of a convex parameter space $\boldsymbol{B}_i \subset \mathbb{R}^p$, where p is the dimension of $\boldsymbol{\beta}_i$;*
(2) $\ln f(S_t|S_{t-1} = 0... = S_{t-r} = 0; \mathbf{x}_t; \boldsymbol{\beta}_i)$ *is concave in $\boldsymbol{\beta}_i$ for all $\{S_t, \mathbf{x}_t\}$ and measurable for all $\boldsymbol{\beta}_i$ in $\boldsymbol{B}_i$;*
(3) *P*$[f(S_t|S_{t-1} = 0... = S_{t-r} = 0; \mathbf{x}_t; \boldsymbol{\beta}_i) \neq f(S_t|S_{t-1} = 0... = S_{t-r} = 0; \mathbf{x}_t; \boldsymbol{\beta}_{i,0})] > 0$ *for all $\boldsymbol{\beta}_i \neq \boldsymbol{\beta}_{i,0}$;*
(4) *E*$[\,|\ln f(S_t|S_{t-1} = 0... = S_{t-i} = 0; \mathbf{x}_t; \boldsymbol{\beta}_i)|\,]$ *exists and is finite for all $\boldsymbol{\beta}_i$ in $\boldsymbol{B}_i$.*

Then, as $T \to \infty$, $\hat{\boldsymbol{\beta}}_i$ exists with probability 1 and $\hat{\boldsymbol{\beta}}_i \xrightarrow{p} \boldsymbol{\beta}_i$.

The first and second order derivatives of the logistic cumulative density function are, $\Lambda(v)' = \Lambda(v) - (1 - \Lambda(v))$ and $\Lambda(v)'' = [1 - 2\Lambda(v)]\Lambda(v)[1 - \Lambda(v)]$, respectively. Thus, the score and Hessian for observation t are, respectively,

$$s(\mathbf{w}_t; \boldsymbol{\beta}_i) = \frac{\partial \ln L}{\partial \boldsymbol{\beta}_i} = [S_t - \Lambda(\mathbf{x}_t' \boldsymbol{\beta}_i)]\mathbf{x}_t; \tag{8}$$

$$\boldsymbol{H}(\mathbf{w}_t; \boldsymbol{\beta}_i) = \frac{\partial s(\mathbf{w}_t; \boldsymbol{\beta}_i)}{\partial \boldsymbol{\beta}_i'} = -\Lambda(\mathbf{x}_t' \boldsymbol{\beta}_i)[1 - \Lambda(x_t' \boldsymbol{\beta}_i)]\mathbf{x}_t \mathbf{x}_t', \tag{9}$$

where $\mathbf{w}_t := (S_t, \mathbf{x}_t')'$. Since $\mathbf{x}_t \mathbf{x}_t'$ is positive definite, $\boldsymbol{H}(\mathbf{w}_t; \boldsymbol{\beta}_i)$ is negative semi-definite and the log-likelihood function is concave, therefore condition (2) of Theorem 1 holds. The last two conditions of Theorem 1 are satisfied under the non-singularity of $\mathrm{E}(\mathbf{x}_t \mathbf{x}_t')$ (see Appendix for details).

Theorem 2. ***Asymptotic normality of conditional MLE***
Let $\mathbf{w}_t = (S_t, \mathbf{x}_t')'$ be jointly stationary and $\hat{\boldsymbol{\beta}}_i \xrightarrow{p} \boldsymbol{\beta}_i$. In addition, consider that

(1) *$\boldsymbol{\beta}_i$ is in the interior of $\boldsymbol{B}_i$ (identification);*
(2) *$f(S_t|S_{t-1} = 0... = S_{t-i} = 0; \mathbf{x}_t; \boldsymbol{\beta}_i)$ is twice continuously differentiable in $\boldsymbol{\beta}_i$ for all $\mathbf{w}_t$;*
(3) *E*$[s(\mathbf{w}_t; \boldsymbol{\beta}_{0,r})] = 0$ *and -E*$[\boldsymbol{H}(\mathbf{w}_t; \boldsymbol{\beta}_r)]$ *=E*$[s(\mathbf{w}_t; \boldsymbol{\beta}_{0,r}) s(\mathbf{w}_t; \boldsymbol{\beta}_{0,r})']$ *where $s(\boldsymbol{w}_t; \boldsymbol{\beta}_i)$ and $\boldsymbol{H}(\boldsymbol{w}_t; \boldsymbol{\beta}_i)$ are as defined in (8) and (9) (local dominance condition on the Hessian);*
(4) *for some neighborhood $\mathcal{N}$ of $\boldsymbol{\beta}_i$,*

$$E[\ \sup_{\boldsymbol{\beta}_i \in B} ||\boldsymbol{H}(\mathbf{w}_t; \boldsymbol{\beta}_i)||\ ] < \infty,$$

so that for any consistent estimator $\tilde{\boldsymbol{\beta}}_i$, $\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{H}(\mathbf{w}_t; \tilde{\boldsymbol{\beta}}_i) \xrightarrow{p} E[\boldsymbol{H}(\mathbf{w}_t; \boldsymbol{\beta}_i)]$;
(5) *E*$[\boldsymbol{H}(\mathbf{w}_t; \boldsymbol{\beta}_i)]$ *is nonsingular.*

Thus, if conditions (1) - (5) hold, $\hat{\boldsymbol{\beta}}_i$ is asymptotically normal with

$$Avar(\hat{\boldsymbol{\beta}}_i) = \left(E[\boldsymbol{H}(\mathbf{w}_t; \boldsymbol{\beta}_i)]\right)^{-1} \boldsymbol{\Sigma}_r \left(E[\mathbf{H}(\mathbf{w}_t; \boldsymbol{\beta}_i)]\right)^{-1}$$

where $\boldsymbol{\Sigma}_i$ is the long-run variance of $\{s(\mathbf{w}_t; \boldsymbol{\beta}_i)\}$.

Assuming that $\hat{\boldsymbol{\Sigma}}_i$ is a consistent estimator of $\boldsymbol{\Sigma}_i$, then a consistent estimator of the asymptotic variance of $\hat{\boldsymbol{\beta}}_i$ is

$$\widehat{Avar}\,(\hat{\boldsymbol{\beta}}_i) = \left\{ \frac{1}{T} \sum_{t=1}^{T} \mathbf{H}(\mathbf{w}_t; \hat{\boldsymbol{\beta}}_i) \right\}^{-1} \hat{\boldsymbol{\Sigma}}_i \ \left\{ \frac{1}{T} \sum_{t=1}^{T} \mathbf{H}(\mathbf{w}_t; \hat{\boldsymbol{\beta}}_i) \right\}^{-1}. \tag{10}$$

Theorem 3. *Considering the results of Theorems 1 and 2 it follows as $T \to \infty$ that,*

$$\sqrt{T}\hat{\boldsymbol{\beta}}_i \xrightarrow{d} N(\boldsymbol{\beta}_i, \mathit{Avar}(\hat{\boldsymbol{\beta}}_i)).$$

Theorem 4. *Let Assumption A2 hold. From application of the Delta method it follows that,*

$$\sqrt{T}\, p_i(\mathbf{x}) \xrightarrow{d} N\Big(\Lambda(\mathbf{x}'\boldsymbol{\beta}_i), [\Lambda(\mathbf{x}'\boldsymbol{\beta}_i)']^2\, \mathbf{x}'\, \mathrm{Avar}\,(\hat{\boldsymbol{\beta}}_i)\, \mathbf{x} \Big).$$

The positive Harris recurrence of $S_t|\mathbf{x}_t$ is crucial to ensure that the process moves from one state to another an infinite number of times as $T \to \infty$. This prevents, for example, from having too many zeros in the sequence of S_t (i.e., that y_t crosses z_1 too few times), which results in inaccurate estimates of $\boldsymbol{\beta}_i$ and $p_i(\mathbf{x}) \to 1$.

2.4. Covariate-dependent ET

The covariate-dependent ET to cross z_1 when the process y_t starts at z_0 is,

$$E(T_{z_1}|\mathbf{x}) = \sum_{t=1}^{\infty} t\ P(T_{z_1} = t|\mathbf{x}). \tag{11}$$

If S_t is a first order Markov Chain, i.e. $r = 1$, then $P(S_t = 0|S_{t-1} = 0; \mathbf{x}) = p_1(\mathbf{x})$ and

$$E(T_{z_1}|\mathbf{x}) = \big[1 - p_1(\mathbf{x})\big] \sum_{t=1}^{\infty} t\, p_1(\mathbf{x})^{t-1} = \Big[1 - p_1(\mathbf{x})\Big]^{-1}.$$

Theorem 5. *Let $\widehat{E(T_{z_1}|\mathbf{x})} = \Big[1 - \widehat{p_1(\mathbf{x})}\Big]^{-1}$. For $r = 1$,*

$$\widehat{E(T_{z_1}|\mathbf{x})} \xrightarrow{p} E(T_{z_1}|\mathbf{x}) \ \ and$$
$$\sqrt{T}\Big(\widehat{E(T_{z_1}|\mathbf{x})} - E(T_{z_1}|\mathbf{x})\Big) \xrightarrow{d} N\Big(0, \big[\mathbf{x}\exp(\mathbf{x}'\boldsymbol{\beta}_1)\big]'\mathrm{Avar}\,\big(\widehat{p_1(\mathbf{x})}\big)\,\big[\mathbf{x}\exp(\mathbf{x}'\boldsymbol{\beta}_1)\big]\Big),$$

where $\mathit{Avar}\,\big(\widehat{p_1(\mathbf{x})}\big) := [\Lambda(\mathbf{x}'\boldsymbol{\beta}_1)']^2\, \mathbf{x}'\, \mathit{Avar}(\hat{\boldsymbol{\beta}}_1)\, \mathbf{x}$, and $0 < p_1 < 1$.

Using (3) and (11) we have for $r > 1$ that

$$E(T_{z_1}|\mathbf{x}) = \sum_{t=1}^{r} t\,\big[1 - p_t(\mathbf{x})\big] + \prod_{j=1}^{t-1} p_j(\mathbf{x}) \left\{ \big[1 - p_r(\mathbf{x})\big] \prod_{j=1}^{r-1} p_j(\mathbf{x}) \right\} \sum_{t=r+1}^{\infty} t\; p_r(\mathbf{x})^{t-r}$$

$$= \sum_{t=1}^{r} t\,\big[1 - p_t(\mathbf{x})\big] + \left\{ \big[1 - p_r(\mathbf{x})\big] \prod_{j=1}^{r-1} p_j(\mathbf{x}) \right\} \frac{p_r(\mathbf{x})[1 + r - r\,p_r(\mathbf{x})]}{[1 - p_r(\mathbf{x})]^2}. \tag{12}$$

By the continuous mapping theorem, if $\hat{\boldsymbol{\beta}}_i$ is consistent then $\widehat{E(T_{z_1}|\mathbf{x})}$ will also be a consistent estimator of $E(T_{z_1}|\mathbf{x})$ since it is a continuous function of $\hat{\boldsymbol{\beta}}_i$.

From (12) it follows that for $r \geq 1$ we have,

$$\begin{aligned}
r = 1 &\Rightarrow E(T_{z_1}|\mathbf{x}) = \frac{1}{1 - p_1(\mathbf{x})}, \\
r = 2 &\Rightarrow E(T_{z_1}|\mathbf{x}) = \frac{1 + p_1(\mathbf{x}) - p_2(\mathbf{x})}{1 - p_2(\mathbf{x})}, \\
r = 3 &\Rightarrow E(T_{z_1}|\mathbf{x}) = \frac{p_1(\mathbf{x})\big(p_2(\mathbf{x}) + 1\big) - p_3(\mathbf{x})\big(p_1(\mathbf{x}) + 1\big) + 1}{1 - p_3(\mathbf{x})}, \text{ etc.}
\end{aligned}$$

Therefore, as already stated, it is critical to have consistent estimates of $\boldsymbol{\beta}_i$, $1 \leq i \leq r$. In small samples, S_t may not move from one state to another a sufficient number of times when r is relatively large and estimating these parameters may be problematic. In practice, the choice of r depends on the sample size, the level of persistence, the starting point z_0 and the threshold z_1. We suggest that r is chosen based on two indicators: the residuals of the regression of y_t on its own r lags and $\mathbf{x}_t$, and the statistical significance of $\boldsymbol{\beta}_r$. It is also possible to estimate r using some information criteria such as BIC (Bayesian information criteria); see, for instance, Katz (1981) and Raftery (1985). However, this approach is cumbersome, since it requires estimating the entire transition probability matrix for several Markov chains of different orders, while we are only interested in the probability in (4).

As is evident from (12), an exact asymptotic expression for the distribution of $\widehat{E(T_{z_1}|\mathbf{x})}$ is difficult to obtain since it is a complex non-linear function of $\hat{\boldsymbol{\beta}}_i$. However, advances in computing have made resampling techniques, in particular bootstraping approaches, a valuable tool for the estimation of standard errors and for the construction of confidence intervals.

In this work, suitable bootstrap methods, which allow for serial dependence, are applied. Many different bootstrap techniques for dependent data have been proposed (see, for instance, MacKinnon 2007, Section 6 for a brief overview). A widely used approach in this context is the block bootstrap algorithm (Härdle *et al.* 2003). The block bootstrap consists in dividing the time series into several blocks of b consecutive observations in order to preserve the original structure within a block, and to re-sample the blocks, which may be overlapping or non-overlapping and of fixed or of variable length, as in e.g. the stationary block bootstrap proposed by Politis and Romano (1994).

Lahiri (2003, Chapter 5) compares the performance of four block bootstrap approaches[7] and shows that, in terms of their MSEs, the overlapping block bootstrap outperforms the non-overlapping and the stationary block bootstrap procedures. This conclusion is valid if the block length increases as the sample size T increases at a rate not slower than the optimal rate $\kappa T^{1/3}$, where κ is constant.

Thus, in what follows we will employ the overlapping block bootstrap, also known as "blocks of blocks" bootstrap, proposed by Politis and Romano (1992b). Defining $\boldsymbol{Z}_t \equiv (y_t, \mathbf{x}_t)$, we construct $T - b + 1$ overlapping blocks as

$$\boldsymbol{Z}_1, ..., \boldsymbol{Z}_b, \ \ \boldsymbol{Z}_2, ..., \boldsymbol{Z}_{b+1}, ..., \boldsymbol{Z}_{T-b+1}, ..., \boldsymbol{Z}_T, \tag{13}$$

which are re-sampled in the usual way, using an iid random variable on $\{1, 2, ..., T - b + 1\}$. The block bootstrap algorithm consists of the following steps:

Step 1: Choose the block length b. In the empirical application, we opt for $b = T^{1/3}$;

Step 2: Resample the blocks as illustrated in (13) and generate the bootstrap sample $(y_t^*, \mathbf{x}_t^*)$;

Step 3: Build the process S_t^* in (1) using y_t^* and estimate the covariate-dependent probabilities in (4) and (5) ;

Step 4: Compute $\widehat{E(T_{z_1}|\mathbf{x})}^*$.

Step 5: Repeat Steps 1 to 4 a B number of times, where B is the number of bootstrap simulations, and compute the empirical distribution of $\widehat{E(T_{z_1}|\mathbf{x})}^*$ and respective confidence intervals.

3. Monte Carlo Analysis

This section investigates the finite sample properties of the parameter estimates $\hat{\boldsymbol{\beta}}_r$. We generate the $S_t|\mathbf{x}_t$ process by simulating two-state Markov chains of orders $r = 1, 2, ..., 5$.

In order to simplify the simulation exercise but without loss of generality we make some simplifying assumptions about the data generation process (DGP) of $S_t|\mathbf{x}_t$. In specific, we assume that $p_r(\mathbf{x})$ is covariate-dependent and the remaining probabilities are constant and equal to 0.5. In practice, one would expect that all transition probabilities depend on covariates. As stated in Section 2.3, these assumptions have no effect on the consistency of $\hat{\boldsymbol{\beta}}_r$ since the part of the log-likelihood which corresponds to $S_{t-1} = 0, ..., S_{t-r} = 0$ is maximized individually.

7. In addition to the overlapping, non-overlapping and stationary block bootstraps, Lahiri (2003) also considers the circular block bootstrap proposed by Politis and Romano (1992a).

As an illustration, consider a second order ($r = 2$) Markov chain. In this case, the transition probabilities matrix is completely defined by the probabilities

- $p_2(\mathbf{x}) = P(S_t = 0|S_{t-1} = 0, S_{t-2} = 0; \mathbf{x}) = \Lambda(\mathbf{x}'\boldsymbol{\beta}_2)$,
- $P(S_t = 0|S_{t-1} = 0, S_{t-2} = 1; \mathbf{x}) = 0.5$,
- $P(S_t = 0|S_{t-1} = 1, S_{t-2} = 0; \mathbf{x}) = 0.5$,
- $P(S_t = 0|S_{t-1} = 1, S_{t-2} = 1; \mathbf{x}) = 0.5$,

and our interest centers exclusively on $p_2(\mathbf{x})$, which can be estimated by maximizing the log-likelihood function in (6) for $i = r = 2$. However, it is noteworthy that, since the DGP is a second order Markov chain, $p_1(\mathbf{x}) = P(S_t|S_{t-1}; \mathbf{x})$, also needed to compute the ET, will depend on the first two probabilities presented above, that is, on $p_2(\mathbf{x})$ and $P(S_t = 0|S_{t-1} = 0, S_{t-2} = 1; \mathbf{x})$.

The DGP also considers that $\mathbf{x}_t := \big(1, x_{2t}\big)'$ and $\boldsymbol{\beta}_2 = \big(\beta_{2,1}, \beta_{2,2}\big)'$ in (4), where $x_{2t} \sim N(0, 1)$. Therefore, as $T \to \infty$,

$$
\begin{aligned}
&\frac{\sum_{t=1}^{T} p_2(\mathbf{x}_t)}{T} \xrightarrow{p} E\big(p_2(\mathbf{x}_t)\big) \\
&= \int_{-\infty}^{+\infty} P(S_t = 0|S_{t-1}, ..., S_{t-r} = 0; \mathbf{x}_t) f(x_{2t}) dx_{2t} \\
&= \int_{-\infty}^{+\infty} \frac{\exp^{\mathbf{x}_t'\boldsymbol{\beta}_2}}{1 + \exp^{\mathbf{x}_t'\boldsymbol{\beta}_2}} \frac{1}{\sqrt{2\pi}} \mathsf{exp}^{-\frac{x_{2t}^2}{2}} dx_{2t}.
\end{aligned}
$$

We will investigate two cases, $\boldsymbol{\beta}_2 = (0.0, 3)'$ and $\boldsymbol{\beta}_2 = (2.3, 3)'$ which imply that $\frac{\sum_{t=1}^{T} p_r(\mathbf{x}_t)}{T} \to 0.5$ and $\frac{\sum_{t=1}^{T} p_r(\mathbf{x}_t)}{T} \to 0.75$, respectively. The second case is particularly relevant since we are interested in $z_1 = \bar{y}$ and the macroeconomic variables tend to exhibit some persistence (or slow mean reversion after a shock), which results in higher values of $E(T_{z_1}|\mathbf{x})$.

As the order of the Markov chain is unknown in practice, for each Markov chain of order r generated in the simulations, we estimate first up to fifth order Markov chains.

Table 1 summarizes the Monte Carlo results for $T \in \{500, 1000, 2000\}$. When $i \geq r$, the $\boldsymbol{\beta}_2$ parameters seem to be consistently estimated even for $T = 500$. As expected, the mean of the parameter estimates is closer to the "true" parameter values and the standard deviation of the estimates reduces in all cases when the sample size increases. Although MLE seems to produce consistent estimates of $\boldsymbol{\beta}_2$ even when Markov chains of higher order than the one considered in the DGP is estimated ($i > r$), these estimates are less accurate since

$$
p_{r-1}(\mathbf{x})p_{r-2}(\mathbf{x})...p_1(\mathbf{x}) = P(S_{t-1} = 0, ..., S_{t-r} = 0|\mathbf{x})
$$

decreases as r increases and there are less cases with $\{S_{t-1} = 0, ..., S_{t-r} = 0\}$. Thus, $\boldsymbol{\beta}_r$ will be estimated using a smaller number of observations, since $\delta_i = 1$ in expression (6) occurs less often.

		$T=500$				$T=1000$				$T=2000$				$T=500$				$T=1000$				$T=2000$			
		$\beta_{i,1}=0.0$		$\beta_{i,2}=3.0$		$\beta_{i,1}=0.0$		$\beta_{i,2}=3.0$		$\beta_{i,1}=0.0$		$\beta_{i,1}=3.0$		$\beta_1=2.3$		$\beta_2=3.0$		$\beta_1=2.3$		$\beta_2=3.0$		$\beta_1=2.3$		$\beta_2=3.0$	
r	i	**mean**	**s.d.**	**mean**	**s.d.**	**mean**	**s.d.**	**mean**	**s.d.**	**mean**	**s.d.**	**mean**	s.d.	**mean**	**s.d.**	**mean**	**s.d.**	**mean**	**s.d.**	**mean**	**s.d.**	**mean**	**s.d.**	**mean**	**s.d.**
	1	-0.005	0.196	3.076	0.407	-0.002	0.137	3.035	0.278	0.001	0.099	3.019	0.193	2.340	0.289	3.067	0.384	2.322	0.201	3.035	0.263	2.312	0.139	3.017	0.185
	2	-0.014	0.284	3.157	0.621	-0.008	0.196	3.077	0.413	-0.001	0.138	3.037	0.281	2.351	0.342	3.090	0.458	2.327	0.236	3.047	0.310	2.314	0.162	3.023	0.216
1	3	-0.032	0.448	3.362	1.043	-0.017	0.283	3.160	0.629	-0.007	0.197	3.076	0.411	2.370	0.409	3.129	0.553	2.335	0.277	3.064	0.365	2.319	0.189	3.031	0.254
	4	-0.054	0.673	3.541	1.391	-0.031	0.434	3.332	0.998	-0.018	0.289	3.158	0.630	2.397	0.499	3.179	0.686	2.345	0.325	3.084	0.429	2.325	0.222	3.041	0.295
	5	0.008	0.799	3.202	1.382	-0.033	0.624	3.353	1.178	-0.036	0.453	3.365	1.114	2.430	0.609	3.247	0.846	2.363	0.387	3.119	0.515	2.334	0.259	3.057	0.346
	1	0.008	0.137	0.800	0.169	0.004	0.098	0.791	0.118	0.003	0.069	0.786	0.083	0.824	0.195	0.993	0.185	0.816	0.136	0.980	0.129	0.812	0.096	0.976	0.091
	2	-0.010	0.287	3.169	0.627	-0.006	0.197	3.079	0.410	-0.003	0.137	3.036	0.279	2.358	0.389	3.109	0.525	2.330	0.264	3.055	0.350	2.317	0.183	3.029	0.243
2	3	-0.023	0.441	3.353	1.030	-0.015	0.288	3.169	0.636	-0.007	0.197	3.075	0.406	2.385	0.470	3.162	0.646	2.341	0.310	3.075	0.412	2.323	0.215	3.039	0.284
	4	-0.039	0.667	3.507	1.371	-0.031	0.448	3.354	1.013	-0.014	0.289	3.164	0.623	2.420	0.580	3.233	0.812	2.358	0.369	3.105	0.487	2.330	0.252	3.052	0.332
	5	0.007	0.812	3.225	1.397	-0.029	0.622	3.358	1.165	-0.035	0.452	3.390	1.138	2.470	0.720	3.331	1.000	2.379	0.446	3.147	0.595	2.338	0.296	3.071	0.391
	1	0.015	0.127	0.366	0.142	0.007	0.091	0.359	0.100	0.005	0.066	0.358	0.071	0.433	0.178	0.492	0.149	0.423	0.126	0.483	0.103	0.422	0.088	0.483	0.073
	2	0.008	0.194	0.815	0.246	0.003	0.136	0.797	0.168	0.003	0.097	0.791	0.118	0.828	0.263	1.005	0.250	0.818	0.186	0.987	0.176	0.816	0.13	0.981	0.123
3	3	-0.026	0.439	3.354	1.037	-0.014	0.287	3.158	0.624	-0.008	0.193	3.071	0.403	2.416	0.566	3.215	0.778	2.357	0.369	3.105	0.488	2.328	0.252	3.049	0.332
	4	-0.039	0.647	3.520	1.380	-0.030	0.440	3.331	0.990	-0.015	0.282	3.158	0.621	2.460	0.689	3.301	0.963	2.383	0.444	3.152	0.591	2.338	0.295	3.068	0.391
	5	0.008	0.802	3.220	1.384	-0.033	0.617	3.342	1.179	-0.039	0.446	3.371	1.095	2.501	0.814	3.389	1.117	2.410	0.541	3.208	0.727	2.348	0.345	3.093	0.462
	1	0.020	0.128	0.186	0.137	0.010	0.089	0.179	0.094	0.006	0.065	0.175	0.066	0.247	0.168	0.263	0.136	0.236	0.119	0.256	0.098	0.228	0.084	0.253	0.068
	2	0.022	0.184	0.381	0.207	0.011	0.127	0.364	0.141	0.005	0.092	0.360	0.099	0.444	0.250	0.507	0.207	0.434	0.175	0.495	0.147	0.423	0.124	0.486	0.101
4	3	0.016	0.281	0.851	0.360	0.010	0.192	0.812	0.243	0.003	0.135	0.799	0.167	0.849	0.378	1.052	0.367	0.833	0.261	1.012	0.247	0.817	0.182	0.989	0.168
	4	-0.042	0.659	3.504	1.369	-0.028	0.441	3.336	0.984	-0.018	0.281	3.165	0.622	2.506	0.838	3.409	1.159	2.414	0.548	3.211	0.740	2.351	0.361	3.101	0.482
	5	-0.007	0.798	3.250	1.402	-0.031	0.620	3.374	1.167	-0.045	0.455	3.377	1.119	2.515	0.934	3.452	1.265	2.444	0.650	3.277	0.881	2.369	0.433	3.140	0.583
	1	0.029	0.127	0.097	0.133	0.016	0.090	0.093	0.092	0.009	0.065	0.091	0.064	0.151	0.156	0.144	0.133	0.136	0.111	0.138	0.094	0.126	0.078	0.133	0.066
	2	0.036	0.179	0.195	0.197	0.021	0.127	0.184	0.135	0.009	0.091	0.179	0.093	0.267	0.231	0.277	0.195	0.247	0.166	0.263	0.135	0.232	0.118	0.256	0.095
5	3	0.042	0.260	0.398	0.305	0.023	0.185	0.374	0.204	0.008	0.129	0.364	0.140	0.471	0.345	0.539	0.301	0.445	0.244	0.507	0.203	0.426	0.174	0.491	0.141
	4	0.026	0.424	0.921	0.609	0.017	0.277	0.844	0.358	0.004	0.192	0.817	0.247	0.894	0.553	1.142	0.567	0.850	0.371	1.048	0.353	0.821	0.253	1.004	0.234
	5	0.004	0.791	3.295	1.404	-0.030	0.601	3.353	1.169	-0.041	0.445	3.363	1.093	2.469	1.004	3.379	1.281	2.477	0.764	3.341	1.007	2.392	0.530	3.200	0.722

Table 1. Means and standard deviations of Markov chain parameter estimates

Notes for Table 1: r refers to the Markov chain order considered in the data generating process and $i = 1,..,5$ is the order of the Markov chain used to estimate $\boldsymbol{\beta}_i$. All results presented are based on 10000 Monte Carlo simulations.

4. Empirical Application

It is nowadays widely accepted that the relation between major economic variables is nonlinear (see, for instance, Terasvirta *et al.* 2010). However, since most existing nonlinear models assume a parametric functional form and require the estimation of a considerable numbers of parameters, alternative ways to capture some aspects of the nonlinear relationships have been developed. For instance, in order to cater for the possibility that the yield curve predicts more accurately when drastic changes in output occur, some authors considered a binary dependent variable which equals one when the National Bureau of Economic Research (NBER) dates a recessions and zero otherwise (see, for instance, Estrella and Hardouvelis 1991 and Estrella and Mishkin 1998). Then, discrete choice models, such as logit or probit, are employed to estimate the effect of YS on the probability of a recession.

Most standard models for binary responses do not take the dynamic structure of the data into account, which is crucial for applications to time series data. A relevant exception is the parametric (linear) dynamic probit model which includes lags of the binary response variable in the probit function; see e.g. Kauppi and Saikkonen (2008) and Antunes *et al.* (2018).

An alternative approach to incorporate information provided by the past values of the dependent variable is to assume the Markovian property. In this context, the interpretation is in terms of transitional rather than marginal probabilities (see, for instance, Azzalini 1994) and the focus is on the estimation of the probabilities of transitions between states. The framework that we introduce in this paper, which also relies on the Markov assumption, allows us to obtain covariate-dependent transition probabilities without requiring a rigid parametric functional form[8] or the estimation of a large number of parameters. Moreover, instead of simply indicating the presence or not of a recession, the binary variable is given by (1). Since the proposed approach allow us to consider different threshold z_1 and starting z_0 values, it may be useful to capture additional information about possible nonlinear relationships between a dependent variable and a set of covariates.

However, the major advantage of our approach is that it provides a simple method to estimate the covariate-dependent expected time (ET) to cross a threshold, which may be a useful reduced-form tool to investigate relevant topics such as dynamic controllability. Roughly speaking, Buiter and Gersovitz (1981) define that a system is dynamically controllable if a path for the economic instruments exists which is capable of moving the vector with the economic objectives from any initial value to any other target value in pre-assigned finite time. They argue that this ability to achieve a vector of target values is relevant for economic policy even if these target values cannot be maintained. Thus, by choosing a target value z_1 and calculating covariate-dependent ET for different

8. The only parametric assumption is that the covariates influence the transition probabilities via a logistic function, as defined in (4) and (5).

starting points z_0 we may gather evidence about the effectiveness of the covariates in driving a dependent variable towards z_1.

This section provides an empirical application to the economic activity-yield spread relationship. The link between YS and economic growth is related to monetary policy, which influences the shape of the yield curve - the representation of several yields or interest rates across different contract lengths - over the business cycles. For instance, monetary policy influences directly, through the use of open market operations, the level of short-term maturity yields. Central banks have a reference interest rate as an important pro-cyclical instrument to pursue their objectives (e.g., price stability). They will lower short run yields in recessions in an attempt to stimulate the economy and will do the opposite when there are inflationary pressures. However, as short-term interest rates have been close to their zero lower bound in recent years, central banks also began to influence long-term interest rates using unconventional monetary policy operations such as, quantitative easing, in order to stimulate aggregate demand and avoid a scenario of low growth and deflation.

Therefore, central banks' monetary policies exert a strong influence on YS. When the economy is in recession, monetary policy actions that have led YS to positive values will promote a faster economic recovery. On the other hand, if the economy is in expansion and the inflation rate suggests that we are facing an over heated economy, monetary policy actions can be taken in order to reduce the YS and slow down economic growth; for instance, by increasing the reference interest rates.

We consider YS as a proxy for the monetary policy stance and investigate how this variable affects the ET that IP growth rate takes to return to its mean after an exogenous shock. In practice, this will be done considering several starting values z_0, each of which correspond to a different S_t process, and by estimating the vector of parameters $\boldsymbol{\beta}_i$ that indicates how YS influences the ET for IP growth rate to return to its mean.

4.1. Data

The proposed methodology is applied to U.S. data. The IP index was selected as an indicator of economic activity due to its higher (monthly) frequency and faster availability relatively to GDP, the most commonly used measure of economic activity. IP seems an adequate choice since, at least for the more industrialized countries, the value added by industrial production represents a substantial share of GDP. Moreover, the IP index exhibits more cyclical fluctuations than the financial index. It is expected that the larger number of observations available and the greater cyclical variability of the IP index will have a positive impact on parameter estimation. We consider the monthly seasonally adjusted U.S. IP index obtained from OECD's Main Economic Indicators Publication, for the period from December 1964 to February 2019 (651 observations).

YS, which is the difference between the long-term[9] and the short-term interest rates[10], used as a proxy for the monetary policy stance, is obtained from OECD's Monthly Monetary and Financial Statistics. Figure 1 graphically presents both time series.

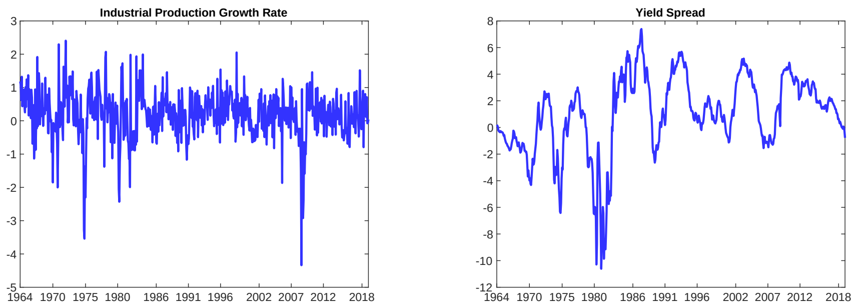


Figure 1: U.S. Industrial Production (monthly) growth rate and yield spread.

4.2. Empirical Results

Let y_t be the monthly IP growth rate and x_t YS. We consider the empirical mean of the IP growth rate as the threshold, $z_1 = \overline{y}$, and estimate $\boldsymbol{\beta}_i$ $(1 \leq i \leq r)$ for the covariate-dependent probabilities defined in (4) and (5) considering several starting values z_0, each corresponding to a different S_t process as defined in (1). The $\boldsymbol{\beta}_i$ parameters are crucial to properly estimate the impact of the covariates on the ET for y_t to cross z_1 when it starts from z_0; see (12).

YS has been identified as an important leading indicator in the literature. For instance, Estrella and Mishkin (1998) conclude that the steepness of the yield curve is an accurate predictor of real activity, especially between two and six quarters ahead. Thus, we estimated the probabilities in (4) and (5) using k-periods lagged YS as covariates, with $k = 1, ..., 6$, considering Markov chains of different orders and starting values z_0. Since, overall the 3-month lagged YS provides stronger statistical evidence, we will use this variable as a covariate. In other words, we will consider that $\mathbf{x}_t = (1, x_{t-3})'$ and $\boldsymbol{\beta}_i = (\beta_{i,1}, \beta_{i,2})$, with $i = 1, ..., r$.

The Markov chain order r was chosen based on the analysis of the residuals of a regression of y_t on its own lags and YS. Additionally, we also take into account the statistical significance of $\beta_{i,1}$ and $\beta_{i,2}$ for several starting values z_0. Thus, a third order Markov process (i.e. $r = 3$) has been considered. For $r > 3$, $\hat{\beta}_{r,2}$ is not

9. Long-term interest rates are computed using government securities with outstanding maturities of 10 years.

10. Short-term interest rates are either the three month interbank rate associated to loans provided and taken among banks for any excess or shortage of liquidity over several months or the rate associated with Treasury bills, certificates of deposit or comparable instruments, each of three month maturity.

statistically significant and the standard error of $\hat{\beta}_{r,1}$ also becomes substantially higher for almost all starting values z_0 considered.

Table 2 shows that the coefficients of YS ($\hat{\beta}_{i2}, i = 1, 2, 3$) are negative when $z_0 < \overline{y}$ and positive when $z_0 > \overline{y}$. For the first case the YS coefficients are statistically significant at the 10% significance level for almost all starting values considered. When $z_0 > z_1$, significance is only observed for $i = 2, 3$ and if the starting values are not particularly large.

		$i=1$				$i=2$				$i=3$			
		$\hat{\beta}_{1,1}$		$\hat{\beta}_{1,2}$		$\hat{\beta}_{2,1}$		$\hat{\beta}_{2,2}$		$\hat{\beta}_{3,1}$		$\hat{\beta}_{3,2}$	
α	z_0	**coef.**	**prob.**	**coef.**	**prob.**	**coef.**	**prob.**	**coef.**	**prob.**	**coef.**	**prob.**	**coef.**	**prob.**
-1.50	-0.886	0.761	0.011	-0.171	0.049	0.957	0.008	-0.176	0.072	0.998	0.013	-0.131	0.162
-1.25	-0.705	0.960	0.000	-0.183	0.018	1.193	0.000	-0.199	0.033	1.019	0.003	-0.225	0.033
-1.00	-0.524	0.755	0.000	-0.167	0.006	0.944	0.000	-0.118	0.068	0.903	0.001	-0.201	0.027
-0.75	-0.343	0.485	0.003	-0.170	0.001	0.811	0.000	-0.137	0.031	0.964	0.000	-0.190	0.033
-0.50	-0.162	0.509	0.000	-0.148	0.001	0.833	0.000	-0.133	0.018	1.026	0.000	-0.191	0.014
-0.25	0.020	0.319	0.007	-0.101	0.000	0.864	0.000	-0.140	0.009	0.996	0.000	-0.211	0.005
0.25	0.382	0.177	0.102	0.102	0.015	0.282	0.071	0.132	0.020	0.231	0.192	0.254	0.009
0.50	0.563	0.448	0.002	0.061	0.125	0.489	0.006	0.152	0.015	0.438	0.045	0.270	0.003
0.75	0.744	0.595	0.000	0.048	0.227	0.633	0.002	0.138	0.050	0.490	0.032	0.209	0.020
1.00	0.925	0.809	0.000	0.073	0.191	0.742	0.001	0.178	0.042	0.812	0.003	0.184	0.063
1.25	1.106	1.051	0.000	0.110	0.154	0.972	0.000	0.167	0.089	1.092	0.001	0.134	0.167
1.50	1.287	0.958	0.000	0.042	0.363	0.955	0.002	0.147	0.157	0.991	0.005	0.125	0.219

Table 2. Estimated parameters for $p_i(\mathbf{x}) := P(S_t = 0|S_{t-1} = 0, ..., S_{t-i} = 0; \mathbf{x})$, $i \leq r$

Notes for Table 2: $z_0 = \overline{y} + \alpha\hat{\sigma}_y$ and $z_1 = \overline{y} \approx 0.201$, where $\overline{y}$ and $\hat{\sigma}_y$ are, respectively, the sample mean and standard deviation.

We computed the unconditional (proposed by Nicolau 2017) and covariate dependent ET time curves[11] (ETC) and their 95% confidence intervals computed using the overlapping block bootstrap described in Subsection 2.4 with 999 bootstrap replications and block length equal to $T^{1/3}$.

Figure 2 shows the estimated (unconditional and covariate dependent) ETCs and their 95% confidence bounds considering 24 different starting points z_0, equally spaced in the interval $\left(\overline{y} - 1.5\hat{\sigma}_y, \overline{y} + 1.5\hat{\sigma}_y\right)$, with $\overline{y}$ and $\hat{\sigma}_y$ the sample mean and standard deviation of y_t, respectively, and six different values for the explanatory variable z_{t-3}.

The bootstrap-based ETC estimates presented in Figure 2 seem to confirm the results in Table 2. That is, when the IP growth rate is below its mean, $z_0 < \bar{y}$, the β_{i2} estimates are negative. As a consequence, the ET to reach $z_1 \geq \bar{y}$ is relatively low if YS is positive. Therefore, the IP growth rate easily recovers from low values when YS is positive; however, recovery is slower when YS is negative. For example,

11. As in Nicolau (2017), we call ET curve to the graphical representation of the ET estimates for different starting values z_0, but same threshold $z_1 = \overline{y}$

consider the case where YS is equal to 3.9% (see Figure 2 panel F). If the initial value of the IP growth rate is negative, recovery is fast and takes around 2 months on average to reach $\bar{y}$. On the other hand, if YS is negative and equal to -1.9%, recovery is much slower and can take on average four months (see Figure 2 panel A).

When the IP growth rate is above its mean, $z_0 > \bar{y}$, the estimates of β_{i2} become positive and the opposite interpretation applies. Thus, in this case, the ET to decrease to $z_1 \leq \bar{y}$ is delayed if YS is positive (prosperity tends to last) and accelerated if YS is negative. For example, if the YS is equal to 3.9% (see Figure 2 panel F), the IP growth rate will remain above its mean value for about 4 months while the same only happens for two months on average if YS is negative and equal to -1.9% (see Figure 2 panel A).

Hence, we can conclude that there is statistical evidence that YS influences the ET the IP growth rate takes to return to its mean, which is graphically reflected in the asymmetric shapes of the conditional ET curves for high and low YS values. For instance, panels A, E and F of Figure 2 illustrate this feature particularly well.

The relationship between YS and economic activity has been investigated by an extensive literature. Harvey (1989), Stock and Watson (1989), Estrella and Hardouvelis (1991) and Estrella and Mishkin (1998), among others, found statistical evidence that YS predicts future output growth. Most of this research is based on a linear framework of analysis (OLS regressions), considering an appropriate lead-lag relationship. However, there are some exceptions to this practice. For instance, Galbraith and Tkacz (2000) used the linearity tests against TAR models suggested by Hansen (1996) and found evidence in support of the asymmetric impact of YS on the conditional expectation of output growth. When YS is above a specific threshold value, the additional effect of a large positive spread becomes small and statistically insignificant.

Moreover, YS has also been successfully used in predicting recessions (see, for instance, Estrella and Hardouvelis 1991, Estrella and Mishkin 1998 and Kauppi and Saikkonen 2008). Most of this evidence was obtained using probit or logit models where the dependent variable used was a recession indicator, which equals 1 when the NBER dates a recession and zero otherwise.

The proposed methodology is somehow related with this strand of the literature since it considers a logit specification for the transition probabilities in (4) and (5). However, it is much more flexible and can provide additional information about the economic activity-yield spread relationship. We considered $z_1 = \bar{y}$ in order to illustrate the proposed approach, but other threshold values such as $z_1 = 0$ could also have been considered. With $z_1 = \bar{y}$, the covariate-dependent ET curve may provide important insights about the predictive content of YS. For instance, let us consider that $y_{t-1} > \bar{y}$. If YS is positively related to future economic activity, a negative YS increases the probability that the dependent variable will cross $\bar{y}$ in period t and consequently ET will be lower.

A visual informal analysis of the ET curves presented in Figure 2 suggests some interesting facts. First, for $z_0 > \bar{y}$, the ET estimates change very little when YS increases from -1.9 to -0.6 (see panels A and B) and from 0 to 1 (see panels C and D). It seems that a large YS value is necessary to increase the ET substantially (see panels E and F). On the other hand, if $z_0 < \bar{y}$, the decrease in the ET estimates when YS increases seems clearer.

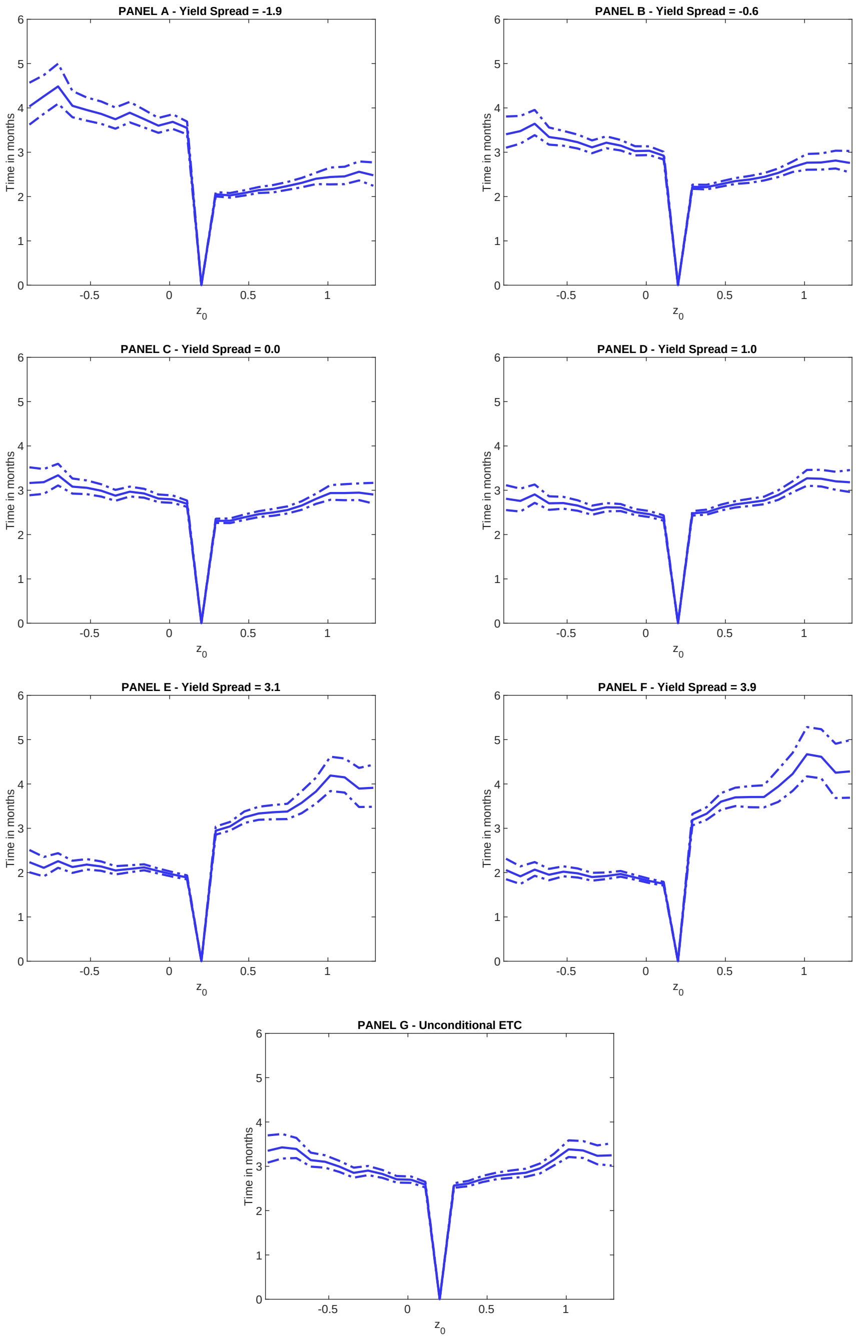


Figure 2: Unconditional and conditional estimated expected time curves (ETC) with overlapping block bootstrap 95% CI. For the conditional case the following six values for the covariate were considered: $(\bar{x}-\hat{\sigma}_x) = -1.9$, $Q_1 = -0.6$, 0 , $\bar{x}$=1 , $Q_3 = 3.1$, $(\bar{x}-\hat{\sigma}_x) = 3.9$; where Q_1 and Q_3 are the first and third quartiles of x_t, respectively.

5. Conclusions

In this paper we propose a simple and easy to implement approach to investigate the effect of covariates on the expected time (ET) to cross a threshold given a specific starting point. In order to estimate the parameters that describe the relationship between the ET and the covariates, we adapt the procedure to estimate Markov models of any order proposed by Islam and Chowdhury (2006). We confirm via Monte Carlo simulations that the relevant parameters, $\boldsymbol{\beta}_r$, are consistently estimated even when the sample size is relatively small $(T = 500)$.

However, since the expression for ET in (12) is a highly nonlinear function of $\boldsymbol{\beta}_i$, with $i = 1, .., r$, we consider an overlapping block-bootstrap procedure to obtain the standard errors of the ET estimates and to construct relevant confidence intervals. Existing literature on the topic suggests that this block-bootstrap variant is a good choice to resample dependent data. We used it to obtain confidence intervals for the ET that the U.S. IP growth rate takes to revert to its mean given a starting point and a particular YS value. Figure 2 shows that the width of the confidence intervals is relatively narrow even when the starting value z_0 is far (in absolute value) from the threshold value z_1 (which results in less accurate estimates of $\boldsymbol{\beta}_i$).

The empirical application to the U.S. economy shows that there is statistical evidence supporting the influence of YS on the ET for the IP growth rate to return to its mean. Namely, a high YS reduces this ET if the IP growth rate starts by assuming below average values. Thus, monetary policies that result in higher YS could play an important role in stimulating a faster return to desirable growth rates in periods of weak growth or contraction. Moreover, the YS value seems also critical when the IP growth rate is larger than average. If YS is negative, the IP growth rate will return quickly to below average values. This finding may be related to the widely documented ability of the yield curve inversion (negative YS) to predict recessions (see, for instance, Estrella and Mishkin 1998).

The application of the proposed methodology to the economic activity-yield spread relationship illustrates that it provides insights that may be relevant to policymakers. Thus, the proposed approach can be adapted to support a wide range of economic decisions since it provides a flexible and easy to implement framework that allows us to infer about the nonlinear relationship between a dependent variable associated with an economic objective and a set of relevant covariates associated with economic policy instruments.

References

Antunes, António, Diana Bonfim, Nuno Monteiro, and Paulo M. M. Rodrigues (2018). "Forecasting banking crises with dynamic panel probit models." *International Journal of Forecasting*, 34(2), 249 – 275.

Audretsch, David B. and Talat Mahmood (1995). "New Firm Survival: New Results Using a Hazard Function." *The Review of Economics and Statistics*, 77(1), 97–103.

Azzalini, A. (1994). "Logistic Regression for Autocorrelated Data with Application to Repeated Measures." *Biometrika*, 81(4), 767–775.

Buiter, Willem H. and Mark Gersovitz (1981). "Issues in controllability and the theory of economic policy." *Journal of Public Economics*, 15(1), 33 – 43.

Cox, D.R. and Valerie Isham (1980). *Point Processes*. Routledge.

Durbin, J. (1971). "Boundary-crossing probabilities for the Brownian motion and Poisson processes and techniques for computing the power of the Kolmogorov-Smirnov test." *Journal of Applied Probability*, 8(3), 431–453.

Engle, Robert F. and Jeffrey R. Russell (1998). "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data." *Econometrica*, 66(5), 1127–1162.

Estrella, Arturo and Gikas A. Hardouvelis (1991). "The Term Structure as a Predictor of Real Economic Activity." *The Journal of Finance*, 46(2), 555–576.

Estrella, Arturo and Frederic S. Mishkin (1998). "Predicting U.S. Recessions: Financial Variables as Leading Indicators." *The Review of Economics and Statistics*, 80(1), 45–61.

Florio, Anna (2004). "The Asymmetric Effects of Monetary Policy." *Journal of Economic Surveys*, 18(3), 409–426.

Galbraith, John W. and Greg Tkacz (2000). "Testing for asymmetry in the link between the yield spread and output in the G-7 countries." *Journal of International Money and Finance*, 19(5), 657 – 672.

Giesecke, Kay (2006). "Default and information." *Journal of Economic Dynamics and Control*, 30(11), 2281 – 2303.

Hansen, Bruce E. (1996). "Inference When a Nuisance Parameter Is Not Identified Under the Null Hypothesis." *Econometrica*, 64(2), 413–430.

Harvey, Campbell R. (1989). "Forecasts of Economic Growth from the Bond and Stock Markets." *Financial Analysts Journal*, 45(5), 38–45.

Hayashi, Fumio (2000). *Econometrics*. Princeton University Press.

Härdle, Wolfgang, Joel Horowitz, and Jens-Peter Kreiss (2003). "Bootstrap Methods for Time Series." *International Statistical Review*, 71(2), 435–459.

Islam, M. Ataharul and Rafiqul Islam Chowdhury (2006). "A higher order Markov model for analyzing covariate dependence." *Applied Mathematical Modelling*, 30(6), 477 – 488.

Katz, Richard W. (1981). "On Some Criteria for Estimating the Order of a Markov Chain." *Technometrics*, 23(3), 243–249.

Kauppi, Heikki and Pentti Saikkonen (2008). "Predicting U.S. Recessions with Dynamic Binary Response Models." *The Review of Economics and Statistics*, 90(4), 777–791.

Kiefer, Nicholas M. (1988). "Economic Duration Data and Hazard Functions." *Journal of Economic Literature*, 26(2), 646–679.

Lahiri, S. N. (2003). *Resampling Methods for Dependent Data*. Springer-Verlag New York.

Lo, Andrew W., A.Craig MacKinlay, and June Zhang (2002). "Econometric models of limit-order executions." *Journal of Financial Economics*, 65(1), 31 – 71.

MacKinnon, James G. (2007). "Bootstrap Hypothesis Testing." Working Paper 1127, Economics Department, Queen's University.

Mata, José and Pedro Portugal (2002). "The survival of new domestic and foreign-owned firms." *Strategic Management Journal*, 23(4), 323–343.

McCall, B. P. (1994). "Testing the proportional hazards assumption in the presence of unmeasured heterogeneity." *Journal of Applied Econometrics*, 9(3), 321–334.

Meyer, Bruce D. (1990). "Unemployment Insurance and Unemployment Spells." *Econometrica*, 58(4), 757–782.

Meyn, Sean, Richard L. Tweedie, and Peter W. Glynn (2009). *Markov Chains and Stochastic Stability*. Cambridge Mathematical Library, 2 ed., Cambridge University Press.

Nicolau, JoÃ£o (2017). "A simple nonparametric method to estimate the expected time to cross a threshold." *Statistics & Probability Letters*, 123, 146 – 152.

Ohn, Jonathan, Larry W. Taylor, and Adrian Pagan (2004). "Testing for duration dependence in economic cycles." *The Econometrics Journal*, 7(2), 528–549.

Politis, D. and J.P. Romano (1992a). *Exploring the Limits of Bootstrap*, chap. A Circular Block Resampling Procedure for Stationary Data, pp. 263–270. Wiley, New York.

Politis, Dimitris N. and Joseph P. Romano (1992b). "A General Resampling Scheme for Triangular Arrays of α-Mixing Random Variables with Application to the Problem of Spectral Density Estimation." *The Annals of Statistics*, 20(4), 1985–2007.

Politis, Dimitris N. and Joseph P. Romano (1994). "The Stationary Bootstrap." *Journal of the American Statistical Association*, 89(428), 1303–1313.

Raftery, Adrian E. (1985). "A Model for High-Order Markov Chains." *Journal of the Royal Statistical Society. Series B (Methodological)*, 47(3), 528–539.

Saldías, Martín (2017). "The Nonlinear Interaction Between Monetary Policy and Financial Stress." *IMF Working Papers*, 17, 1.

Sichel, Daniel E. (1991). "Business Cycle Duration Dependence: A Parametric Approach." *The Review of Economics and Statistics*, 73(2), 254–260.

Stock, James H. and Mark W. Watson (1989). "New Indexes of Coincident and Leading Economic Indicators." *NBER Macroeconomics Annual*, 4, 351–394.

Sueyoshi, Glenn T. (1995). "A Class of Binary Response Models for Grouped Duration Data." *Journal of Applied Econometrics*, 10(4), 411–431.

Terasvirta, Timo, Dag Tjostheim, and Clive Granger (2010). *Modelling Nonlinear Economic Time Series*. Oxford University Press.

Whitmore, G. A. (1986). "First-Passage-Time Models for Duration Data: Regression Structures and Competing Risks." *Journal of the Royal Statistical Society. Series D (The Statistician)*, 35(2), 207–219.

Appendix

The Markov chain's log likelihood function

For an rth order Markov chain, the log-likelihood function can be expressed as the sum of r^2 components. As an illustration, consider a second order Markov chain and define:

$\delta_{000} = 1$ if $\{S_t = 0, S_{t-1} = 0, S_{t-2} = 0\}$;
$\delta_{100} = 1$ if $\{S_t = 0, S_{t-1} = 0, S_{t-2} = 1\}$;
$\delta_{010} = 1$ if $\{S_t = 0, S_{t-1} = 1, S_{t-2} = 0\}$;
$\delta_{110} = 1$ if $\{S_t = 0, S_{t-1} = 1, S_{t-2} = 1\}$;
$p_{000}(\mathbf{x}) := p_2(\mathbf{x}) = P(S_t = 0|S_{t-1} = 0, S_{t-1} = 0)$;
$p_{100}(\mathbf{x}) = P(S_t = 0|S_{t-1} = 0, S_{t-1} = 1)$;
$p_{010}(\mathbf{x}) = P(S_t = 0|S_{t-1} = 1, S_{t-1} = 0)$;
$p_{110}(\mathbf{x}) = P(S_t = 0|S_{t-1} = 1, S_{t-1} = 1)$.

The log-likelihood for observation t can be expressed as

$$\ln L = L_1 + L_2 + L_3 + L_4,$$

where

$$L_1 = \delta_{001} \ln(1 - p_{000}(\mathbf{x})) + \delta_{000} \ln(p_{000}(\mathbf{x})),$$
$$L_2 = \delta_{101} \ln(1 - p_{100}(\mathbf{x})) + \delta_{100} \ln(p_{100}(\mathbf{x})),$$
$$L_3 = \delta_{011} \ln(1 - p_{010}(\mathbf{x})) + \delta_{010} \ln(p_{010}(\mathbf{x})),$$
$$L_4 = \delta_{111} \ln(1 - p_{110}(\mathbf{x})) + \delta_{110} \ln(p_{110}(\mathbf{x})).$$

Proof of Proposition 1

Consider the probability $P(S_t = 0|S_{t-1} = 0, S_{t-2} = 0, ..., S_0 = 0; \mathbf{x})$. The results for other cases are similar. The event $\{S_{t-1} = 1, S_{t-2} = 1, ..., S_0 = 1\}$ represents $\{y_{t-1} < z_1, y_{t-2} < z_1, ..., y_1 < z_1, y_0 \leq z_0\}$. Therefore,

$$P(S_t = 1|S_{t-1} = 1, S_{t-2} = 1, ..., S_0 = 1; \mathbf{x})$$
$$\equiv P(y_t < z_1|y_{t-1} < z_1, y_{t-2} < z_1, ..., y_0 \leq z_0; \mathbf{x})$$

and since y_t is an rth order Markov process,

$$P(S_t = 1|S_{t-1} = 1, S_{t-2} = 1, ..., S_0 = 1; \mathbf{x})$$
$$= P(S_t = 1|S_{t-1} = 1, S_{t-2} = 1, ..., S_{t-r} = 1; \mathbf{x})$$
$$= P(y_t < z_1|y_{t-1} < z_1, y_{t-2} < z_1, ..., y_{t-r+1} \leq z_1, y_{t-r} \leq z_0; \mathbf{x}).$$

■

Proof of Theorem 1

For condition (3) (conditional density identification) note that

$$E[(\mathbf{x}_t\boldsymbol{\beta}_{i,0} - \mathbf{x}_t\boldsymbol{\beta}_i)^2] = E[\{\mathbf{x}_t(\boldsymbol{\beta}_{i,0} - \boldsymbol{\beta}_i)\}^2] = (\boldsymbol{\beta}_{i,0} - \boldsymbol{\beta}_i)'E(\mathbf{x}_t\mathbf{x}_t')(\boldsymbol{\beta}_{i,0} - \boldsymbol{\beta}_i) > 0\ ,$$

where $\boldsymbol{\beta}_i$ is the true parameter vector and $\boldsymbol{\beta}_{i,0}$ a parameter vector such that $\boldsymbol{\beta}_{i,0} \neq \boldsymbol{\beta}_i$. Hence, $\mathbf{x}_t\boldsymbol{\beta}_i \neq \mathbf{x}_t\boldsymbol{\beta}_{i,0}$ with positive probability and since $\Lambda(v)$ is strictly monotonic, we have $\Lambda(\mathbf{x}_t\boldsymbol{\beta}_i) \neq \Lambda(\mathbf{x}_t\boldsymbol{\beta}_{i,0})$ when $\mathbf{x}_t\boldsymbol{\beta}_i \neq \mathbf{x}_t\boldsymbol{\beta}_{i,0}$.

Condition (4) holds if $E[|\log(f(S_t|S_{t-1} = 0... = S_{t-r} = 0; \mathbf{x}_t; \boldsymbol{\beta}_i)|] < \infty$ for all $\boldsymbol{\beta}_i$.

For the logistic function, it is easy to verify that

$$|\ln \Lambda(v)| \leq |\ln \Lambda(0)| + |v|.$$

Furthermore, note that

$$\begin{aligned}
|\ln f(S_t|S_{t-1} = 0, ..., S_{t-r} = 0; \mathbf{x}_t; \boldsymbol{\beta}_i)| &\leq |S_t|\ \ln(\Lambda(\mathbf{x}_t\boldsymbol{\beta}_i))| + |1 - S_t|\ |\ln(1 - \Lambda(\mathbf{x}_t\boldsymbol{\beta}_i))| \\
&\leq |\ln(\Lambda(\mathbf{x}_t\boldsymbol{\beta}_i))| + |\ln(\Lambda(-\mathbf{x}_t\boldsymbol{\beta}_i))| \\
&\text{(since } |S_t| \leq\ 1 \text{ and } |1 - S_t| \leq 1) \\
&\leq 2[\,|\ln(\Lambda(0)| + ||\mathbf{x}_t|| \times ||\boldsymbol{\beta}_i||\,] \\
&\text{(due to the Cauchy-Schwartz inequality)}
\end{aligned}$$

The nonsingularity of $\mathsf{E}(\mathbf{x}_t\mathbf{x}_t')$ implies $\mathsf{E}(x_{it}^2) < \infty$ for all i and, therefore, $\mathsf{E}(||\mathbf{x}_t^2||) < \infty$ and $\mathsf{E}(||\mathbf{x}_t||) < \infty$. Thus, the nonsingularity of $\mathsf{E}(\mathbf{x}_t\mathbf{x}_t')$ ensures that the logit ML estimator is consistent. ∎

Proof of Theorem 2

Condition (1) is satisfied for the logit model if the compact parameter space $\boldsymbol{B}_i$ is taken to be $\mathbb{R}^p$. Condition (2) is obviously satisfied. To check condition (3) note that since $\mathsf{E}[S_t|S_{t-1} = 0, ..., S_{t-r} = 0; \mathbf{x}_t] = \Lambda(\mathbf{x}_t\boldsymbol{\beta}_i)$, we have $\mathsf{E}[s(\mathbf{w}_t; \boldsymbol{\beta}_i)|\mathbf{x}_t] = 0$ and, by the Law of Total Expectations $\mathsf{E}[s(\mathbf{w}_t; \boldsymbol{\beta}_i)] = 0$.

In order to derive the conditional information matrix, note that, using the standard rules of differentiation we have that,

$$\mathsf{E}\left[\frac{\partial \ln L}{\partial\boldsymbol{\beta}_i\partial\boldsymbol{\beta}_i'}\right] = -\mathsf{E}\left[\frac{\partial \ln L}{\partial\boldsymbol{\beta}_i}\ \frac{\partial \ln L}{\partial\boldsymbol{\beta}_i'}\right] + \mathsf{E}\left[\frac{1}{\mathsf{ln}L}\ \frac{\partial^2 \mathsf{ln}L}{\partial\boldsymbol{\beta}_i\partial\boldsymbol{\beta}_i'}\right],$$

where it is easy to verify that the second term is zero. Thus, the following relationship between the expected value of the Hessian matrix and the expected outer product of the scores holds:

$$-\mathsf{E}[\boldsymbol{H}(\mathbf{w}_t; \boldsymbol{\beta}_i)] = \mathsf{E}[\mathbf{s}(\mathbf{w}_t; \boldsymbol{\beta}_i)\ \mathbf{s}(\mathbf{w}_t; \boldsymbol{\beta}_i)'],$$

where $s(\mathbf{w}_t; \boldsymbol{\beta}_i)$ and $\boldsymbol{H}(\mathbf{w}_t; \boldsymbol{\beta}_i)$ are the functions defined in (8) and (9),respectively. Regarding the local dominance of the Hessian - condition (4) -, since $\Lambda(\mathbf{x}_t'\boldsymbol{\beta}_i)[1 -$

$\Lambda(\mathbf{x}_t'\boldsymbol{\beta}_i)] < 1$, we have $||\boldsymbol{H}(\mathbf{w}_t;\boldsymbol{\beta}_i)|| \leq ||\mathbf{x}_t\mathbf{x}_t'||$ for all $\boldsymbol{\beta}_i$. It can be shown that $\mathsf{E}[||\mathbf{x}_t\mathbf{x}_t'||] < \infty$ if $\mathsf{E}[\mathbf{x}_t\mathbf{x}_t']$ is nonsingular (and hence finite). Finally, condition (5) also requires that $\mathsf{E}[\mathbf{x}_t\mathbf{x}_t']$ is nonsingular. ■

Proof of Theorem 4

Under assumption A2, the joint stationarity of $\{y_t, \mathbf{x}_t\}$ implies the joint stationarity of $\{S_t, \mathbf{x}_t\}$, given the measurability of (1). ■

Working Papers

2017

1|17 The diffusion of knowledge via managers' mobility
Giordano Mion | Luca David Opromolla | Alessandro Sforza

2|17 Upward nominal wage rigidity
Paulo Guimarães | Fernando Martins | Pedro Portugal

3|17 Zooming the ins and outs of the U.S. unemployment
Pedro Portugal | António Rua

4|17 Labor market imperfections and the firm's wage setting policy
Sónia Félix | Pedro Portugal

5|17 International banking and cross-border effects of regulation: lessons from Portugal
Diana Bonfim | Sónia Costa

6|17 Disentangling the channels from birthdate to educational attainment
Luís Martins | Manuel Coutinho Pereira

7|17 Who's who in global value chains? A weighted network approach
João Amador | Sónia Cabral | Rossana Mastrandrea | Franco Ruzzenenti

8|17 Lending relationships and the real economy: evidence in the context of the euro area sovereign debt crisis
Luciana Barbosa

9|17 Impact of uncertainty measures on the Portuguese economy
Cristina Manteu | Sara Serra

10|17 Modelling currency demand in a small open economy within a monetary union
António Rua

11|17 Boom, slump, sudden stops, recovery, and policy options. Portugal and the Euro
Olivier Blanchard | Pedro Portugal

12|17 Inefficiency distribution of the European Banking System
João Oliveira

13|17 Banks' liquidity management and systemic risk
Luca G. Deidda | Ettore Panetti

14|17 Entrepreneurial risk and diversification through trade
Federico Esposito

15|17 The portuguese post-2008 period: a narrative from an estimated DSGE model
Paulo Júlio | José R. Maria

16|17 A theory of government bailouts in a heterogeneous banking system
Filomena Garcia | Ettore Panetti

17|17 Goods and factor market integration: a quantitative assessment of the EU enlargement
FLorenzo Caliendo | Luca David Opromolla | Fernando Parro | Alessandro Sforza

2018

1|18 Calibration and the estimation of macroeconomic models
Nikolay Iskrev

2|18 Are asset price data informative about news shocks? A DSGE perspective
Nikolay Iskrev

3|18 Sub-optimality of the friedman rule with distorting taxes
Bernardino Adão | André C. Silva

4|18 The effect of firm cash holdings on monetary policy
Bernardino Adão | André C. Silva

5|18 The returns to schooling unveiled
Ana Rute Cardoso | Paulo Guimarães | Pedro Portugal | Hugo Reis

6|18 Real effects of financial distress: the role of heterogeneity
Francisco Buera | Sudipto Karmakar

7|18 Did recent reforms facilitate EU labour market adjustment? Firm level evidence
Mario Izquierdo | Theodora Kosma | Ana Lamo | Fernando Martins | Simon Savsek

8|18 Flexible wage components as a source of wage adaptability to shocks: evidence from European firms, 2010–2013
Jan Babecký | Clémence Berson | Ludmila Fadejeva | Ana Lamo | Petra Marotzke | Fernando Martins | Pawel Strzelecki

9|18 The effects of official and unofficial information on tax compliance
Filomena Garcia | Luca David Opromolla Andrea Vezulli | Rafael Marques

10|18 International trade in services: evidence for portuguese firms
João Amador | Sónia Cabral | Birgitte Ringstad

11|18 Fear the walking dead: zombie firms, spillovers and exit barriers
Ana Fontoura Gouveia | Christian Osterhold

12|18 Collateral Damage? Labour Market Effects of Competing with China – at Home and Abroad
Sónia Cabral | Pedro S. Martins | João Pereira dos Santos | Mariana Tavares

13|18 An integrated financial amplifier: The role of defaulted loans and occasionally binding constraints in output fluctuations
Paulo Júlio | José R. Maria

14|18 Structural Changes in the Duration of Bull Markets and Business Cycle Dynamics
João Cruz | João Nicolau | Paulo M.M. Rodrigues

15|18 Cross-border spillovers of monetary policy: what changes during a financial crisis?
Luciana Barbosa | Diana Bonfim | Sónia Costa | Mary Everett

16|18 When losses turn into loans: the cost of undercapitalized banks
Laura Blattner | Luísa Farinha | Francisca Rebelo

17|18 Testing the fractionally integrated hypothesis using M estimation: With an application to stock market volatility
Matei Demetrescu | Paulo M. M. Rodrigues | Antonio Rubia

2019

1|19 The transmission of unconventional monetary policy to bank credit supply: evidence from the TLTRO
António Afonso | Joana Sousa-Leite

2|19 How responsive are wages to demand within the firm? Evidence from idiosyncratic export demand shocks
Andrew Garin | Filipe Silvério

3|19 Vocational high school graduate wage gap: the role of cognitive skills and firms
Joop Hartog | Pedro Raposo | Hugo Reis

4|19 What is the Impact of Increased Business Competition?
Sónia Félix | Chiara Maggi

5|19 Modelling the Demand for Euro Banknotes
António Rua

6|19 Testing for Episodic Predictability in Stock Returns
Matei Demetrescu | Iliyan Georgiev
Paulo M. M. Rodrigues | A. M. Robert Taylor

7|19 The new ESCB methodology for the calculation of cyclically adjusted budget balances: an application to the Portuguese case
Cláudia Braz | Maria Manuel Campos
Sharmin Sazedj

8|19 Into the heterogeneities in the Portuguese labour market: an empirical assessment
Fernando Martins | Domingos Seward

9|19 A reexamination of inflation persistence dynamics in OECD countries: A new approach
Gabriel Zsurkis | João Nicolau | Paulo M. M. Rodrigues

10|19 Euro area fiscal policy changes: stylised features of the past two decades
Cláudia Braz | Nicolas Carnots

11|19 The Neutrality of Nominal Rates: How Long is the Long Run?
João Valle e Azevedo | João Ritto | Pedro Teles

12|19 Testing for breaks in the cointegrating relationship: on the stability of government bond markets' equilibrium
Paulo M. M. Rodrigues | Philipp Sibbertsen
Michelle Voges

13|19 Monthly Forecasting of GDP with Mixed Frequency MultivariateSingular Spectrum Analysis
Hossein Hassani | António Rua | Emmanuel Sirimal Silva | Dimitrios Thomakos

14|19 ECB, BoE and Fed Monetary-Policy announcements: price and volume effects on European securities markets
Eurico Ferreira | Ana Paula Serra

15|19 The financial channels of labor rigidities: evidence from Portugal
Edoardo M. Acabbi | Ettore Panetti | Alessandro Sforza

16|19 Sovereign exposures in the Portuguese banking system: determinants and dynamics
Maria Manuel Campos | Ana Rita Mateus | Álvaro Pina

17|19 Time vs. Risk Preferences, Bank Liquidity Provision and Financial Fragility
Ettore Panetti

18|19 Trends and cycles under changing economic conditions
Cláudia Duarte | José R. Maria | Sharmin Sazedj

19|19 Bank funding and the survival of start-ups
Luísa Farinha | Sónia Félix | João A. C. Santos

20|19 From micro to macro: a note on the analysis of aggregate productivity dynamics using firm-level data
Daniel A. Dias | Carlos Robalo Marques

21|19 Tighter credit and consumer bankruptcy insurance
António Antunes | Tiago Cavalcanti | Caterina Mendicino | Marcel Peruffo | Anne Villamil

2020

1|20 On-site inspecting zombie lending
Diana Bonfim | Geraldo Cerqueiro | Hans Degryse | Steven Ongena

2|20 Labor earnings dynamics in a developing economy with a large informal sector
Diego B. P. Gomes | Felipe S. Iachan | Cezar Santos

3|20 Endogenous growth and monetary policy: how do interest-rate feedback rules shape nominal and real transitional dynamics?
Pedro Mazeda Gil | Gustavo Iglésias

4|20 Types of International Traders and the Network of Capital Participations
João Amador | Sónia Cabral | Birgitte Ringstad

5|20 Forecasting tourism with targeted predictors in a data-rich environment
Nuno Lourenço | Carlos Melo Gouveia António Rua

6|20 The expected time to cross a threshold and its determinants: A simple and flexible framework
Gabriel Zsurkis | João Nicolau | Paulo M. M. Rodrigues