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

Gabriel Zsurkis\textsuperscript{a}, João Nicolau\textsuperscript{b} and Paulo M. M. Rodrigues\textsuperscript{c}

\textsuperscript{a} Banco de Portugal and ISEG-Universidade de Lisboa
\textsuperscript{b} ISEG-Universidade de Lisboa and CEMAPRE
\textsuperscript{c} Banco de Portugal and Nova School of Business and Economics

November 29, 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 analysis of the response of U.S. unemployment persistence to monetary policy and government spending shocks is provided, contributing to a still limited literature which simultaneously allows for both types of shocks. Our results suggest that unexpected monetary and fiscal expansions seem to have a relevant role in accelerating the pace of unemployment decline towards its natural rate; and that contractionary monetary and fiscal shocks in a context of labor market slack may result in high ETs.

Keywords: Expected time, Markov chains, nonlinearity, unemployment gap, natural rate of unemployment, fiscal shocks, monetary shocks.

JEL classification: C32, C41, C51, E63
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 economics, finance, biology and the 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.

The extant literature on duration analysis 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. 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); in the analysis of firm survival (see, for instance, Audretsch and Mahmood; 1995 and Mata and Portugal; 2002); and in the analysis of duration dependence in economic cycles (see, for instance, Sichel; 1991, Ohn et al.; 2004 and Berge and Pfajfar; 2019).

A closely related approach to duration dependence modeling is to treat the occurrence of a given event as a random variable which follows a point process\(^2\). A point process is a sequence of non-negative random variables representing the times at which events occur, which is defined as \(\{t_i\}_{i \in \{1,2,\ldots\}}\), with \(0 \leq t_i \leq t_{i+1}\). A complete description of such processes is formulated in terms of the conditional intensity function which can, roughly speaking, be associated to the probability per unit of time of observing an event in the next

\(^1\)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).
Thus, different parameterizations of this function result in different point process models. Existing models can be grouped into two classes. The first class, 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 occurrences 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.

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 \( (t_1) \). First hitting time problems have been mostly addressed in a continuous time context invoking Wiener processes, which involve complex mathematical concepts and can result in models which are difficult to estimate. Nicolau (2017) introduced an intuitive and easy to implement framework for estimating the first passage time probability function in a discrete time context. One of the main contributions of the present work is the development of a novel approach to estimate transition probabilities allowing for covariates, which corresponds to an important extension of this framework. To this end, we generalize 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 response 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 covariates in driving the outcome variable towards some preassigned values. In practice, the impact of covariates may not be symmetric and may 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 rapid economic growth causes inflationary pressures and an easy one in a recession context in order to boost a rapid economic recovery. However, it has been shown in the literature that the responsiveness of the real economy to monetary shocks is different during recessions and expansions (see, for instance, Romer and Romer; 1994, Florio; 2004, Lo and Piger; 2005 and references therein). The framework introduced in this paper

---

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.
allows us to investigate these possible nonlinear dynamics by estimating ET conditional on different starting values. If, for a given starting value, changes in covariates are reflected in changes in the ET estimates, it suggests that the chosen covariates affect the movement towards a specified threshold in that specific situation. When other starting values are considered conclusions may however differ.

More specifically, the proposed methodology is also a relevant contribution to the vast literature on the effects of fiscal and monetary shocks on economic fluctuations (see Ramey; 2016 for a recent survey). It provides a useful complement to the traditional approach which is based on estimating macroeconomic shock multipliers by providing further information on the sensitivity of an economic variable to macroeconomic policy changes. Rather than measuring the average effect of the shocks during a given time horizon, the interest lies in inferring whether macroeconomic policy shocks promote a faster economic stabilization.

To further illustrate the usefulness of our approach an application to U.S. unemployment persistence is provided. Specifically, we analyse the response of U.S. unemployment persistence to monetary policy and government spending shocks. This is an important contribution, since with a few exceptions, such as e.g. Rossi and Zubairy (2011), most existing literature only focuses on one of these shocks at a time (see, inter alia, Romer and Romer; 2004, Blanchard and Perotti; 2002, Ramey and Zubairy; 2018, and Brinca et al.; 2019). The framework introduced in this paper allows us to consider both monetary and fiscal shocks simultaneously, which provides relevant insights on the impact of the interaction between these two macroeconomic shocks on the dynamics of the unemployment gap.

As a measure of persistence, we consider the expected time the unemployment rate takes to return to its natural rate, where the labor market is in a sustainable equilibrium. Our results suggest that fiscal shocks are more effective than monetary shocks in stimulating a faster return of unemployment to the natural rate of unemployment in a context of labor market slack. These results are in line with recent findings pointing to the importance of fiscal policy for short-run economic stabilization (see, for instance, Romer; 2012). However, the effects of fiscal policy are highly regime dependent, for instance, Auerbach and Gorodnichenko (2012) show that the estimated fiscal multipliers of government purchases are larger in recessions. In our analysis, we consider positive starting values, i.e., $z_0 > 0$ ($z_0 < 0$ was not addressed due to its limited economic relevance), which corresponds to a positive unemployment gap, which is typically associated with negative output gaps and recessions. Thus, indirectly, our focus centers on whether macroeconomic shocks stimulate a faster recovery from economic downturns. Our results also suggest that the relative effect of monetary policy on the ET is lower when $z_0 = 2$ than when $z_0 < 2$. As $z_0 = 2$ suggests that we may be facing a persistent weak demand environment, these findings are in line with the limited effectiveness of expansionary monetary policy in the
presence of bad expectations about the future reported, for instance, in Florio (2004) and Tenreyro and Thwaites (2016).

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 ET and the covariates. Section 4 presents an empirical application which investigates the response of U.S. unemployment persistence to monetary policy and government spending shocks. Section 5 concludes and a Technical Appendix collects detailed proofs of the results presented in the paper.

2 The proposed methodology

2.1 The process and probabilities of interest

Let $x_t$ be a $k \times 1$ vector of covariates and $\{(y_t, x_t)\}$ a vector of discrete-time processes with state space $\mathbb{R}^{k+1}$, characterized by the following Assumption.

**Assumption A.**

(A1) $y_t | x_t$ is a Markov process of order $r$;

(A2) $\{(y_t, x_t)\}$ 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 $\sigma$-finite measure $m(y)$ such that $m(A) > 0$ implies $E(T_A | X_0 = a) < \infty$ for every $a \in D \setminus \overline{A}$, where $\overline{A}$ is the closure of set $A$. Assumption (A2) ensures that $\{y_t\}$ 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, e.g., Meyn et al.; 2009, Chap. 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\(^5\). 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:

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

(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).

\(^5\)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.
Then, the probability that \( y_t \) crosses the threshold \( z_1 \) for the first time, starting from \( z_0 \), is

\[
P(T_{z_1} = t) = P(S_t = 1, S_{t-1} = 0, S_{t-2} = 0, \ldots, S_1 = 0 | S_0 = 0),
\]

which is equivalent to

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

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

**Proposition 1** Considering that Assumption (A1) holds and that \( y_t | x_t \) is a Markov process of order \( r \), then \( S_t | x_t \) is an \( r \)th order two-state Markov chain.

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

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

where \( p_i(x) = P(S_i = 0 | S_{i-1} = 0, \ldots, S_{t-i} = 0 | x) \) for \( 1 \leq i \leq r \).

### 2.2 Covariate-dependent transition probabilities

Considering that Assumption (A1) and Proposition 1 hold, we can treat \( S_t | x_t \) 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

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

with

\[
p_1(x) := \pi_{00}(x) = P(S_t = 0 | S_{t-1} = 0; x) = \frac{\exp x' \beta_1}{1 + \exp x' \beta_1} =: \Lambda(x' \beta_1)
\]

and \( \Lambda(x' \beta_1) \) 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 \( r \)th 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 \( r \)th order Markov chain \( S_t \), i.e.,

\[
\begin{array}{cccc}
m & S_{t-r} & S_{t-(r-1)} & ... & S_{t-1} \\
1 & 0 & 0 & ... & 0 \\
2 & 0 & 0 & ... & 1 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
2^r - 1 & 1 & 1 & ... & 0 \\
2^r & 1 & 1 & ... & 1 \\
\end{array}
\]

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 \( p_1(x), ..., p_r(x) \) are
needed to obtain the probability function for \( T_{z1|x} \). Hence, if \( S_t|x \) is an \( r \)th order Markov
chain, we define,

\[
p_r(x) = P(S_t = 0|S_{t-1} = 0, ..., S_{t-r} = 0; x) = \Lambda(x'\beta_r).
\] (4)

Moreover, for \( j = 1, ..., r - 1 \),

\[
p_j(x) = P(S_t = 0|S_{t-1} = 0..., S_{t-j} = 0; x) = \Lambda(x'\beta_j).
\] (5)

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

Despite similarities between (4) and (5) and the standard logit model for binary re-
sponses, the proposed approach is less restrictive. Firstly, the focus here is on the transi-
tion probabilities between states and not on the conditional probability of success. More-
over, 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 \( i \)th 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
Appendix for details. Thus, we can maximize individually the part of the log-likelihood
function which corresponds to \( S_{t-1}=S_{t-2}=\ldots=S_{t-i}=0 \), considering for observation \( t \) that,

\[
\ln L_i = \ln f(S_t|S_{t-1} = 0, \ldots, S_{t-i} = 0; x_t; \beta_i) \\
= \delta_i \left[ S_i \ln \left( 1 - \Lambda(x'_i \beta_i) \right) + (1 - S_i) \ln \left( \Lambda(x'_i \beta_i) \right) \right],
\]

where \( f(\cdot) \) is a conditional density function, and \( \delta_i \) an indicator variable which is equal to one when \( S_{t-1} = 0, \ldots, S_{t-i} = 0 \) 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 \( \beta_i \), \( 1 \leq i \leq r \), it is crucial that consistent estimates of these coefficient vectors are obtained. In what follows Theorem 1 establishes the consistency results of the estimates and Theorem 2 the normality of the conditional quasi-maximum likelihood estimator (QMLE) of \( \beta_i \).

**Theorem 1 - Consistency of the conditional QMLE without compactness**

Let \( \{S_t, x_t\} \) be jointly stationary with conditional density \( f(S_t|S_{t-1} = \ldots = S_{t-i} = 0; x_t; \beta_i) \) and

\[
\hat{\beta}_i = \arg \max_{\beta \in B_i} \frac{1}{T} \sum_{t=1}^{T} \ln f(S_t|S_{t-1} = \ldots = S_{t-i} = 0; x_t; \beta_i)
\]

is the conditional QMLE. Moreover, consider that,

1. the true parameter vector \( \beta_i \) is an element of the interior of a convex parameter space \( B_i \subset \mathbb{R}^p \), where \( p \) is the dimension of \( \beta_i \);
2. \( \ln f(S_t|S_{t-1} = 0\ldots = S_{t-r} = 0; x_t; \beta_i) \) is concave in \( \beta_i \) for all \( \{S_t, x_t\} \) and measurable for all \( \beta_i \) in \( B_i \);
3. \( P[f(S_t|S_{t-1} = 0\ldots = S_{t-r} = 0; x_t; \beta_i) \neq f(S_t|S_{t-1} = 0\ldots = S_{t-r} = 0; x_t; \beta_{i,0})] > 0 \) for all \( \beta_i \neq \beta_{i,0} \);
4. \( E[|\ln f(S_t|S_{t-1} = 0\ldots = S_{t-r} = 0; x_t; \beta_i)|] \) exists and is finite for all \( \beta_i \) in \( B_i \).

Then, as \( T \to \infty \), \( \hat{\beta}_i \) exists with probability 1 and \( \hat{\beta}_i \overset{p}{\to} \beta_i \).

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

\[
s(w_t; \beta_i) = \frac{\partial \ln L}{\partial \beta_i} = [S_i - \Lambda(x'_i \beta_i)]x_t;
\]

\[
H(w_t; \beta_i) = \frac{\partial^2 \ln L}{\partial \beta_i \partial \beta'_i} = \sum_{j=1}^{T} [S_j - \Lambda(x'_j \beta_i)]x_j x'_j.
\]
where \( w_t := (S_t, x_t') \). Since \( x_t x_t' \) is positive definite, \( H(w_t; \beta_i) \) is negative semi-definite and the log-likelihood function is concave, therefore condition (2) of Theorem 1 holds. Moreover, conditions (3) and (4) of Theorem 1 are satisfied under the non-singularity of \( E(x_t x_t') \) (see Appendix for details).

**Theorem 2 - Asymptotic normality of the conditional QMLE**

Let \( w_t = (S_t, x_t') \) be jointly stationary and \( \hat{\beta}_i \xrightarrow{p} \beta_i \). In addition, consider that

1. \( \beta_i \) is in the interior of \( B_i \) (identification);
2. \( f(S_t|S_{t-1} = 0, \ldots, S_{t-i} = 0; x_t; \beta_i) \) is twice continuously differentiable in \( \beta_i \) for all \( w_t \);
3. \( E[s(w_t; \beta_{0,r})] = 0 \) and \( -E[H(w_t; \beta_r)] = E[s(w_t; \beta_{0,r}) s(w_t; \beta_{0,r})'] \) where \( s(w_t; \beta_i) \) and \( H(w_t; \beta_i) \) are as defined in (8) and (9), respectively (local dominance condition on the Hessian);
4. for some neighborhood \( \mathcal{N} \) of \( \beta_i \),
   \[
   E[\sup_{\beta_i \in B} ||H(w_t; \beta_i)||] < \infty,
   \]
   so that for any consistent estimator \( \hat{\beta}_i \),
   \[
   \frac{1}{T} \sum_{t=1}^{T} H(w_t; \hat{\beta}_i) \xrightarrow{p} E[H(w_t; \beta_i)];
   \]
5. \( E[H(w_t; \beta_i)] \) is nonsingular.

Thus, if conditions (1) - (5) hold, it follows as \( T \to \infty \) that,

\[
\sqrt{T} \hat{\beta}_i \xrightarrow{d} N(\beta_i, Avar(\hat{\beta}_i)).
\]

with

\[
Avar(\hat{\beta}_i) = \left( E[H(w_t; \beta_i)] \right)^{-1} \Sigma_i \left( E[H(w_t; \beta_i)] \right)^{-1}
\]

where \( \Sigma_i \) is the long-run variance of \( \{s(w_t; \beta_i)\} \).

Assuming that \( \hat{\Sigma}_i \) is a consistent estimator of \( \Sigma_i \), then a consistent estimator of the asymptotic variance of \( \hat{\beta}_i \) is

\[
\hat{Avar}(\hat{\beta}_i) = \left\{ \frac{1}{T} \sum_{t=1}^{T} H(w_t; \hat{\beta}_i) \right\}^{-1} \hat{\Sigma}_i \left\{ \frac{1}{T} \sum_{t=1}^{T} H(w_t; \hat{\beta}_i) \right\}^{-1}. \tag{10}
\]

Moreover, Theorem 3 provides the limit distribution of the covariate-dependent transition probabilities \( p_i(x) \), a result which is of importance to characterize the normality of the covariate-dependent ET.
Theorem 3 Let Assumption (A2) hold. From the application of the Delta method it follows as \( T \to \infty \) that,

\[
\sqrt{T} \, p_i(x) \xrightarrow{d} N \left( \Lambda(x' \beta_i), [\Lambda(x' \beta_i)]^2 \, \text{Var} \left( \hat{\beta}_i \right) x \right).
\]

The positive Harris recurrence of \( S_t|x_t \) 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, the sequence \( S_t \) from having too many zeros (i.e., that \( y_t \) crosses \( z_1 \) too few times), avoiding in this way inaccurate estimates of \( \beta_i \) and \( p_i(x) \to 1 \).

2.4 Covariate-dependent ET

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

\[
E(T_{z_1}|x) = \sum_{t=1}^{\infty} t \, P(T_{z_1} = t|x).
\]  

(11)

If \( S_t \) is a first order (\( r = 1 \)) Markov chain, then \( P(S_t = 0|S_{t-1} = 0;x) = p_1(x) \) and

\[
E(T_{z_1}|x) = \left[ 1 - p_1(x) \right] \sum_{t=1}^{\infty} t \, p_1(x)^{t-1} = \left[ 1 - p_1(x) \right]^{-1}.
\]

Hence, the following theorem can be stated.

Theorem 4 Let \( E(T_{z_1}|x) = \left[ 1 - p_1(x) \right]^{-1} \). For \( r = 1 \),

\[
E(T_{z_1}|x) \xrightarrow{p} E(T_{z_1}|x)
\]

and

\[
\sqrt{T} \left( E(T_{z_1}|x) - E(T_{z_1}|x) \right) \xrightarrow{d} N \left( 0, \left[ \Lambda(x' \beta_1) \right]' \text{Var} \left( \hat{\beta}_1 \right) \left[ \Lambda(x' \beta_1) \right] \right),
\]

where \( \text{Var} \left( \hat{p}_1(x) \right) := [\Lambda(x' \beta_1)]^2 x' \text{Var} \left( \hat{\beta}_1 \right) x \), and \( 0 < p_1 < 1 \).

Note that using (3) and (11), for \( r > 1 \) we have that

\[
E(T_{z_1}|x) = \sum_{t=1}^{r} t \left[ 1 - p_t(x) \right] \prod_{j=0}^{t-1} p_j(x) + \left\{ \left[ 1 - p_r(x) \right] \prod_{j=0}^{r-1} p_j(x) \right\} \sum_{t=r+1}^{\infty} t \, p_r(x)^{t-r}
\]

\[
= \sum_{t=1}^{r} t \left[ 1 - p_t(x) \right] \prod_{j=0}^{t-1} p_j(x) + \frac{p_r(x)[1 + r - r \, p_r(x)] \prod_{j=0}^{r-1} p_j(x)}{1 - p_r(x)}
\]

(12)
with $p_0(x) = 1$.

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

It is therefore critical to have consistent estimates of $\beta_i$, $1 \leq i \leq r$. In small samples, $S_i$ 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 $x_t$, and the statistical significance of $\beta_r$. 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 estimation of 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 $E(T_{z_1}|x)$ is difficult to obtain since it is a complex non-linear function of $\hat{\beta}_i$. However, advances in computing have made resampling techniques, in particular bootstrap 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 of 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 variable length, as in e.g. the stationary block bootstrap proposed by Politis and Romano (1994).

Lahiri (2003, Chap. 5) compares the performance of four block bootstrap approaches 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 $\kappa$ is constant. Thus in what follows we will employ the overlapping block bootstrap, also known as the “blocks of blocks” bootstrap, proposed by Politis and Romano (1992a).
Defining $Z_t := (y_t, x_t)$, we construct $T - b + 1$ overlapping blocks as

$$(Z_1, ..., Z_b), (Z_2, ..., Z_{b+1}), ..., (Z_{T-b+1}, ..., Z_T),$$

which are re-sampled 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 $b$;
Step 2: Resample the blocks as illustrated in (13) and generate the bootstrap sample $(y_t^*, x_t^*)$;
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 $\widehat{E}(T_{z_1}|x)$;
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 $\widehat{E}(T_{z_1}|x)$ and respective confidence intervals.

3 Monte Carlo Analysis

This section investigates the finite sample properties of the parameter estimates $\hat{\beta}_r$. We generate $S_t|x_t$ 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 $S_t|x_t$. In specific, we assume that $p_r(x)$ is covariate-dependent and that the remaining probabilities are constant and equal to 0.5. In practice, we expect all transition probabilities to depend on covariates. As stated in Section 2.3, these assumptions have no effect on the consistency of $\hat{\beta}_r$ since the part of the log-likelihood which corresponds to $S_{t-1} = 0, ..., S_{t-r} = 0$ is maximized individually.

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

$$p_2(x) = P(S_t = 0|S_{t-1} = 0, S_{t-2} = 0; x) = \Lambda(x'\beta_2);$$
$$P(S_t = 0|S_{t-1} = 0, S_{t-2} = 1; x) = 0.5;$$
$$P(S_t = 0|S_{t-1} = 1, S_{t-2} = 0; x) = 0.5;$$
$$P(S_t = 0|S_{t-1} = 1, S_{t-2} = 1; x) = 0.5.$$

Our interest centers exclusively on $p_2(x)$, 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, $p_1(x) = P(S_t|S_{t-1}; x)$, which is also needed to compute the ET, will depend on the first two probabilities presented above, that is, on $p_2(x)$ and
The DGP is as in (4) with $x_t := (1, x_{2t})'$ and $\beta_2 := (\beta_1, \beta_2)'$, where $x_{2t} \sim N(0, 1)$. Therefore, as $T \to \infty$,

$$\frac{\sum_{t=1}^T p_2(x_t)}{T} \to E(p_2(x_t)) = \int_{-\infty}^{+\infty} P(S_t = 0 | S_{t-1}, \ldots, S_{t-r} = 0; x_t) f(x_{2t}) dx_{2t} = \int_{-\infty}^{+\infty} \frac{\exp^{x_t \beta_2}}{1 + \exp^{x_t \beta_2}} \frac{1}{\sqrt{2\pi}} \exp^{-x_{2t}^2/2} dx_{2t}.$$  

In terms of the values of $\beta_2$ we will consider two cases in our Monte Carlo experiments: Case A) $\beta_2 = (0.0, 3)'$ and Case B) $\beta_2 = (2.3, 3)'$, which imply that $\sum_{t=1}^T p_r(x_t) \to 0.5$ and $\sum_{t=1}^T p_r(x_t) \to 0.75$, respectively. The second case is particularly relevant since we are interested in $z_1 = \bar{y}$ and macroeconomic variables tend to exhibit some persistence (or slow mean reversion after a shock), which results in larger values of $E(T_{z_1} | x)$. As the order of the Markov chain is unknown in practice, for each Markov chain of order $r$ generated in the simulations we estimate Markov chains of orders one to five.

Table 1 summarizes the Monte Carlo results for $T \in \{500, 1000, 2000\}$. When the order $i$ used to estimate the Markov chain is greater or equal to $r$, $i \geq r$, the $\beta_2$ parameters are consistently estimated. 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 QMLE produces consistent estimates of $\beta_2$ even when Markov chains of orders higher than the one considered in the DGP are estimated ($i > r$), these estimates are less accurate since

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

becomes smaller as $r$ increases and there are less cases with $\{S_{t-1} = 0, \ldots, S_{t-r} = 0\}$, and consequently $\beta_r$ will be estimated using a smaller number of observations, since $\delta_i = 1$ in expression (6) occurs less often.
Table 1: Average and standard deviations of Markov chain parameter estimates

<table>
<thead>
<tr>
<th>Case A</th>
<th>T = 500</th>
<th>T = 1000</th>
<th>T = 2000</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \beta_1 = 0 )</td>
<td>( \beta_2 = 3.0 )</td>
<td>( \beta_1 = 0 )</td>
</tr>
<tr>
<td>( r )</td>
<td>average</td>
<td>s.d.</td>
<td>average</td>
</tr>
<tr>
<td>1</td>
<td>0.005</td>
<td>0.196</td>
<td>0.306</td>
</tr>
<tr>
<td>2</td>
<td>-0.014</td>
<td>0.284</td>
<td>0.317</td>
</tr>
<tr>
<td>3</td>
<td>-0.032</td>
<td>0.448</td>
<td>0.362</td>
</tr>
<tr>
<td>4</td>
<td>-0.054</td>
<td>0.673</td>
<td>0.541</td>
</tr>
<tr>
<td>5</td>
<td>0.008</td>
<td>0.799</td>
<td>2.020</td>
</tr>
<tr>
<td>1</td>
<td>0.008</td>
<td>0.137</td>
<td>0.800</td>
</tr>
<tr>
<td>2</td>
<td>-0.010</td>
<td>0.287</td>
<td>0.319</td>
</tr>
<tr>
<td>3</td>
<td>-0.023</td>
<td>0.441</td>
<td>0.353</td>
</tr>
<tr>
<td>4</td>
<td>-0.009</td>
<td>0.667</td>
<td>0.507</td>
</tr>
<tr>
<td>5</td>
<td>0.007</td>
<td>0.812</td>
<td>2.225</td>
</tr>
<tr>
<td>1</td>
<td>0.015</td>
<td>0.127</td>
<td>0.366</td>
</tr>
<tr>
<td>2</td>
<td>0.008</td>
<td>0.194</td>
<td>0.815</td>
</tr>
<tr>
<td>3</td>
<td>-0.026</td>
<td>0.439</td>
<td>0.354</td>
</tr>
<tr>
<td>4</td>
<td>-1.009</td>
<td>0.672</td>
<td>0.437</td>
</tr>
<tr>
<td>5</td>
<td>0.008</td>
<td>0.802</td>
<td>2.220</td>
</tr>
<tr>
<td>1</td>
<td>0.020</td>
<td>0.128</td>
<td>0.186</td>
</tr>
<tr>
<td>2</td>
<td>0.022</td>
<td>0.184</td>
<td>0.381</td>
</tr>
<tr>
<td>3</td>
<td>0.016</td>
<td>0.281</td>
<td>0.851</td>
</tr>
<tr>
<td>4</td>
<td>-0.002</td>
<td>0.529</td>
<td>0.394</td>
</tr>
<tr>
<td>5</td>
<td>-0.007</td>
<td>0.798</td>
<td>2.520</td>
</tr>
<tr>
<td>1</td>
<td>0.029</td>
<td>0.127</td>
<td>0.097</td>
</tr>
<tr>
<td>2</td>
<td>0.036</td>
<td>0.179</td>
<td>0.195</td>
</tr>
<tr>
<td>3</td>
<td>0.042</td>
<td>0.260</td>
<td>0.398</td>
</tr>
<tr>
<td>4</td>
<td>0.026</td>
<td>0.424</td>
<td>0.921</td>
</tr>
<tr>
<td>5</td>
<td>0.004</td>
<td>0.791</td>
<td>3.295</td>
</tr>
</tbody>
</table>

Note: \( r \) refers to the Markov chain order considered in the DGP and \( i = 1, \ldots, 5 \) is the order of the Markov chain used to estimate \( \beta_i \). All results presented are based on 10000 Monte Carlo simulations.
4 Empirical Application

Achieving maximum sustainable employment as fast as possible is a primary macroeconomic objective. In order to meet this goal, monetary and fiscal policies can be considered. Beyond the traditional interest rate channels, monetary policy also affects economic activity through, for instance, bank credit, asset prices and expectations (Mishkin; 1996). Although there is substantial evidence that monetary policy plays a key role in macroeconomic stabilization, recent works show that its impact on the real economy is not symmetric across the business cycle. However, there is no consensus about the exact form of the phase-dependent dynamics. For instance, Romer and Romer (1994) and Lo and Piger (2005) argue that monetary policy is especially effective during recessions, while Florio (2004) and Tenreyro and Thwaites (2016) suggest that an easy monetary policy is less powerful during economic contractions. Analyzing fiscal policy repercussions on the other hand, is even harder due to the multitude of tax and spending instruments capable of generating a wide range of macroeconomic and distributional effects (Leeper; 2010, Brinca et al.; 2016, Brinca et al.; 2019 and Brinca et al.; 2020). Notwithstanding, Rendhal (2016) shows that the efficacy of fiscal policy is more pronounced in a context of persistently low interest rates and high unemployment, where an increase in government spending increases output and, consequently, lowers the unemployment rate. It is important to recall that there are interactions between monetary and fiscal policies (see Friedman; 1948 and Hilbers; 2005). Therefore, in order to quantify the sensitivity of an economic variable to monetary and fiscal policies, it is necessary to separate endogenous from exogenous changes in policy instruments. Thus, the question of interest then consists of the determination of how economic variables react to unanticipated structural disturbances, commonly called shocks.

A large number of methods have been developed to identify macroeconomic shocks; see e.g. Ramey (2016) for a recent survey. A widely employed framework for identification, due to Sims (1980), is to use structural vector autoregressive (SVARs) models. In a seminal paper, Blanchard and Perotti (2002) used this approach to isolate fiscal shocks and study their effects on aggregate output. Regarding monetary policy shocks, since there is agreement that they mainly refer to surprise changes in interest rates, it is common to extract them from Taylor-type rules (Taylor; 1993, 1999). The use of simple reference rules may provide more robustness than choosing a macroeconomic model across a range of plausible models (see, for instance, Taylor and Williams; 2010). Moreover, as stated by Levin (2014), “simple monetary policy rules can serve as valuable benchmarks in determining the course of monetary policy and explaining those judgments to the public” (p.62).

7 The goals of maximum employment and stable prices are often referred to as the Fed’s “dual mandate” in conducting monetary policy.
Aiming to illustrate the potential of the methodology introduced in this paper, we investigate the importance of monetary and fiscal policy shocks in explaining the U.S. unemployment’s ET to return to its socially efficient equilibrium. Thus, we analyze the impact of these shocks on the dynamics of the unemployment gap, i.e., on the difference between the current rate of unemployment and the (short-term) natural rate of unemployment. Blanchflower and Levin (2015) suggest a framework that includes participation and underemployment gaps in addition to the unemployment gap. Nevertheless, the latter component explains most of the labor market slack.

The conventional view dividing unemployment fluctuations between short-run movements (influenced by monetary policy and other determinants of aggregate demand) and the long-run trend or natural rate (only determined by labor market frictions and unaffected by aggregate demand) has been the dominant paradigm and is frequently embodied in models used by central banks and macroeconomists. However, it is not free of criticism. For instance, Ball (1999) found that monetary policy applied during the recessions of the early 1980s explains a substantial part of how much of the cyclical increase in unemployment became structural in several OECD countries. Blanchard (2003, p.4) seems to agree with these findings when he states that monetary policy affects both the actual and the natural rate of unemployment. The analysis by Coibion et al. (2013) on the contribution of the stance of monetary and fiscal policies to unemployment in the last three recessions also points to this conclusion.

The significantly slower rate of recovery of the unemployment rate towards the natural rate after the end of the 2007-2009 financial crisis has rekindled interest about the causes of the rising persistence of U.S. unemployment leading to the revival of the notion of hysteresis\(^8\) (see, for instance, Summers; 2014). Naturally, if the unemployment gap is mostly explained by structural changes in the labor market\(^9\), macroeconomic shocks have a limited impact in countering these dynamics. One compelling example is the increased job polarization in the U.S. labor market (see, for instance, Autor et al.; 2006). Jaimovich and Siu (2020) show that unemployment persistence is mainly due to routine manual workers in the U.S., who have seen their jobs disappear due to technology and outsourcing in recent years, especially during times of recession. These jobs have been replaced by technology or foreign workers, and it is likely that they are not coming back, causing possibly longer unemployment persistence. On the other hand, it is also possible that high unemployment gaps were caused by adverse macroeconomic policy shocks.

We provide additional insights regarding this issue by using the methodology introduced in Section 2 to estimate the covariate-dependent ET to cross a threshold given different starting values, which can capture nonlinear relationships. Thus, our approach

\(^8\)The theory introduced by Blanchard and Summers (1987) that argues that changes in the natural rate of unemployment can be path-dependent.

\(^9\)Coibion et al. (2013) conclude that monetary and fiscal policies can account only for part of the evolving unemployment persistence.
allows us to infer the effectiveness of macroeconomic policy instruments in fostering a rapid return of the unemployment gap to zero. This analysis will be performed by estimating the expected time to cross the natural rate of unemployment starting from some relevant unemployment gap values, conditional on fiscal and monetary policy shocks. The proposed approach implies a very flexible empirical specification, which is a relevant advantage in view of the recent findings that estimated dynamic responses to monetary policy shocks may not be robust among different identification schemes and empirical specifications; see, for instance, Coibion (2012), Ramey (2016) and Miranda-Agrippino and Ricco (2017). Moreover, with few exceptions, such as Rossi and Zubairy (2011), most existing literature only focuses on one shock at a time. The framework introduced in this paper allows us to simultaneously consider both monetary and fiscal shocks, which may provide important insights on the impact of the interaction between these two macroeconomic shocks on the dynamics of the unemployment gap.

4.1 Data

The unemployment rate (source code: LNS14000000; quarterly aggregation method: average), $U_t$, obtained from the US Bureau of Labor Statistics and the (short-term) natural rate of unemployment $^{10}$, $UN_t$, from the US Congressional Budget Office (CBO), are used to measure the amount of current and projected slack in the labor markets. Thus, we will use the unemployment gap, $UG_t = U_t - UN_t$, as our reference variable. The fiscal (government spending) Blanchard-Perotti shocks, $FS_t$, are obtained from Ramey and Zubairy (2018), and the monetary shocks, $MS_t$, are based on deviations from the well-known Taylor (1993) rule. In order to calculate the monetary policy rule, we use the effective federal funds rate (quarterly aggregation method: average), real potential gross domestic product, real gross domestic product and the implicit price deflator, all retrieved from FRED, Federal Reserve Bank of St. Louis. Although unemployment data is available since 1949, the federal funds rate needed to calculate the monetary shocks is only available after 1954. Hence, our period of analysis is from 1954q4 to 2015q4 (245 observations). Figure 1 graphically presents the short-term natural rate of unemployment, the unemployment gap, and the monetary and fiscal policy shocks, respectively, used in our analysis.

---

$^{10}$The difference between the US Congressional Budget Office’s long and short-term natural rate is fully explained by the incorporation of structural factors that temporarily boosted the natural rate beginning in 2008.
Figure 1: Quarterly U.S. unemployment and macroeconomic policy shocks from 1954q4 to 2015q4.

4.2 Empirical Results

In our analysis, we consider the unemployment gap, $UG_t$, as our reference variable, a threshold value equal to zero ($z_1 = 0$) and some positive starting values $z_0$ in order to evaluate the impact of demand-side macroeconomic shocks on the ET unemployment takes to cross the natural rate of unemployment. As explanatory variables, we consider the contemporaneous fiscal shocks and the monetary shocks lagged four quarters. The choice of the lag order of the monetary shocks follows from available empirical evidence which indicates that monetary policy changes impact economic activity with some delay (see, for instance, Goodhart; 2001). In addition, we also include the natural rate of unemployment lagged one-period as a regressor in order to infer whether the impact of the macroeconomic shocks on the unemployment gap depends on the degree of labor under-utilization. The estimates of the $\beta_i$ parameters of the covariate-dependent probabilities defined in (4) and (5) are crucial to properly estimate the impact of the covariates on the ET for $UG_t$ to cross $z_1 = 0$ when it starts at $z_0$; see (12).

The value of the log-likelihood and the statistical significance of the parameter estimates suggest that a first order Markov chain (i.e. $r = 1$) is a suitable choice to model the probability in (5). We computed the covariate dependent ET time curves and their 95% confidence intervals using the overlapping block bootstrap described in subsection
2.4 with 999 bootstrap replications and blocks of 24 observations, in order to mimic the high persistence of the unemployment rate in the original series. Recall that the binary process $S_t$ is initialized when $UG_t$ crosses $z_0$ given $UG_{t-1} \geq z_0$ (see (1) for details). Moreover, since for $z_0 > 2$ and $z_0 < 0$ this event occurs very few times in the sample, it is not possible to accurately estimate the relevant parameters for these cases. Thus, our analysis will just focus on the starting values $z_0 = 1$ and $z_0 = 2$.

Table 2: Bootstrap-based estimates for $p_1(x) := P(S_t = 0 | S_{t-1} = 0; x)$.

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\beta}_{1,1}$</th>
<th>95% CI</th>
<th>$\hat{\beta}_{1,2}$</th>
<th>95% CI</th>
<th>$\hat{\beta}_{1,3}$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.454</td>
<td>0.445</td>
<td>-0.122</td>
<td>-0.154</td>
<td>0.214</td>
<td>0.170</td>
</tr>
<tr>
<td>2</td>
<td>0.486</td>
<td>0.472</td>
<td>-0.120</td>
<td>-0.173</td>
<td>0.143</td>
<td>0.063</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\beta}_{1,1}$</th>
<th>95% CI</th>
<th>$\hat{\beta}_{1,2}$</th>
<th>95% CI</th>
<th>$\hat{\beta}_{1,3}$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.572</td>
<td>2.523</td>
<td>2.625</td>
<td>0.119</td>
<td>0.150</td>
<td>0.088</td>
</tr>
<tr>
<td>2</td>
<td>2.790</td>
<td>2.712</td>
<td>2.877</td>
<td>-0.121</td>
<td>-0.175</td>
<td>-0.070</td>
</tr>
</tbody>
</table>

Table 2 presents the estimates of the $\beta_i$ parameters when $z_0 = 1$ and $z_0 = 2$. Monetary and fiscal policy shocks have the expected sign in both cases and the estimates are all statistically significant at the 5% significance level. A positive government spending shock decreases the probability that unemployment will remain above the natural rate, while a positive monetary shock (a contractionary shock) will have the opposite effect. An alternative specification in which the lagged natural rate of unemployment is replaced by a constant results in very similar estimates for the monetary and fiscal policy shocks; see Table 2.
Figure 2: Conditional estimated ET curves for $z_0 = 2$, with overlapping block bootstrap 95% confidence intervals. For the baseline case in Panel A the fiscal and monetary shocks equal their sample medians. An expansionary (contractionary) fiscal shock refers to the 0.8 (0.2) sample quantile and an expansionary (contractionary) monetary shock refers to the 0.2 (0.8) sample quantile.

Figure 2 shows the estimated ET curves for $z_0 = 2$ and several values of the covariates. The four scenarios in Panels B - E will be compared with the baseline scenario in Panel A, where fiscal and monetary policy shocks are “neutral” (equal to their sample medians). Panel B shows that if a government spending shock equals its 0.8 sample quantile and the monetary shock equals its sample median, the ET that the unemployment gap takes to return to zero decreases by around 23-24% relatively to the baseline scenario in Panel A. A comparable expansionary monetary policy shock (equal to its 0.2 sample quantile) together with a fiscal shock equal to its sample median, decreases the ET by around 6% (see Panel C) when compared with the baseline scenario. Synchronized fiscal and mon-
etary expansionary shocks, that is, a fiscal shock equal to its 0.8 sample quantile and a monetary shock equal to its 0.2 sample quantile, originates a decrease of ET of approximately 28-29% (see Panel D). On the other hand, Panel E shows that contractionary fiscal and monetary shocks (fiscal and monetary shocks equal to their 0.2 and 0.8 sample quantiles, respectively) may result in very high ETs for unemployment to return to its natural rate (corresponding to an increase of approximately 39-40% when compared with the baseline scenario in Panel A). Similar conclusions can be drawn by analyzing the case when \( z_0 = 1 \) (see Figure A.1 in the appendix).

In sum, results suggest that fiscal shocks are more effective than monetary shocks in stimulating a faster return of unemployment to its natural rate in a context of labor market slack. These results are in line with recent findings pointing to the importance of fiscal policy for short-run economic stabilization (see, for instance, Romer; 2012). However, the effects of fiscal policy may be highly regime dependent. For instance, Auerbach and Gorodnichenko (2012) show, using the direct projections approach of Jordà (2005), that the estimated fiscal multipliers of government purchases are larger in recessions. Since we consider \( z_0 > 0 \) (\( z_0 < 0 \) was not addressed due to its limited economic relevance), our analysis considers a context where the unemployment gap is positive, which is typically associated with negative output gaps and recessions. Thus, indirectly, our focus is on whether macroeconomic shocks stimulate a faster recovery from economic downturns. It is noteworthy from our results that the relative effect of monetary policy on the ET when \( z_0 = 2 \) is lower than when \( z_0 = 1 \) is considered. As the unemployment gap is very persistent (see Figure 1), \( z_0 = 2 \) suggests a persistent weak demand environment. These findings are in line with the lower monetary policies’ effectiveness to move the economy out of a recession reported in, for instance, Florio (2004) and Tenreyro and Thwaites (2016).

As a robustness check of our results, we have redone the empirical analysis by considering two cases: i) we used Romer and Romer (2004)’s updated monetary policy shocks’ series up to 2008Q4 from Coibion et al. (2017) together with the fiscal shocks’ data of Blanchard-Perotti obtained from Ramey and Zubairy (2018) (see Table A.1 in the appendix); and ii) we used the narrative military (spending) news shocks in addition to the Blanchard-Perotti shocks, both obtained from Ramey and Zubairy (2018), together with the monetary shocks used in the main text as well as Romer and Romer (2004)’s shocks (see Table A.2 in the appendix). The results obtained in these two cases show that the fiscal and monetary shocks have the expected sign, that the estimates obtained are all statistically significant at the 5% level and confirm our previous conclusions.
5 Conclusions

In this paper we propose a simple and easy to implement approach to investigate the impact 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 ET and the covariates, we generalize 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, \( \beta_r \), are consistently estimated even when the sample size is relatively small (\( T = 500 \)). However, since the expression for ET in (12) is a nonlinear function of \( \beta_i \), with \( i = 1, \ldots, r \), 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 this approach in our empirical analysis to compute confidence intervals for the ET that the U.S. unemployment rate takes to revert to its natural rate given a specific starting point and different values for the monetary and fiscal policy shocks.

Our analysis reveals that unexpected monetary and fiscal expansions seem to have a relevant role in accelerating the pace of unemployment decline towards its natural rate. On the other hand, contractionary monetary and fiscal shocks in a context of labor market slack (positive unemployment gap) may result in extremely high ETs. Overall, the ET for the unemployment gap to return to zero can be considerably high when the natural rate of unemployment is high, even in a context of expansionary macroeconomic shocks. More precisely, ET is more than two times higher when UN_{t-1} assumes values close to its maximum than when it assumes very low values. Therefore, although results suggest that expansionary monetary and fiscal policies are potentially effective in decreasing the persistence of the unemployment gap, demand expansions should be complemented with labor market structural reforms aiming to reduce the natural rate of unemployment in order to stimulate faster returns to “equilibrium” after recessions.

The application of the proposed methodology to the relationship between unemployment and demand-side macroeconomic policies illustrates that it may provide relevant insights to policymakers. 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.
Acknowledgments

We are grateful to an anonymous referee and Co-Editor, Prof. Juan Francisco Rubio-Ramirez, for their helpful and constructive comments, as well as to Pedro Brinca and Miguel de Faria e Castro for very useful and insightful suggestions. 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.

References


URL: https://economics.mit.edu/files/731


Appendix

The Markov chain’s log likelihood function

For an $r$th order Markov chain, the log-likelihood function can be expressed as the sum of $r^2$ components. As an illustration, consider a second order ($r = 2$) 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 $\{S_t = 0, S_{t-1} = 1, S_{t-2} = 1\}$;

and

\[
p_{000}(x) = p_2(x) = P(S_t = 0|S_{t-1} = 0, S_{t-2} = 0);
p_{100}(x) = P(S_t = 0|S_{t-1} = 0, S_{t-2} = 1);
p_{010}(x) = P(S_t = 0|S_{t-1} = 1, S_{t-2} = 0);
p_{110}(x) = P(S_t = 0|S_{t-1} = 1, S_{t-2} = 1).
\]

The log-likelihood for observation $t$ can be expressed as

\[
\ln L = L_1 + L_2 + L_3 + L_4,
\]

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

Proof of Proposition 1

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

\[
P(S_t = 0|S_{t-1} = 0, S_{t-2} = 0, ..., S_0 = 0; x) \equiv P(y_t < z_1|y_{t-1} < z_1, y_{t-2} < z_1, ..., y_0 \leq z_0; x)
\]

and since $y_t$ is an $r$th order Markov process,

\[
P(S_t = 0|S_{t-1} = 0, S_{t-2} = 0, ..., S_0 = 0; x) = P(S_t = 0|S_{t-1} = 0, S_{t-2} = 0, ..., S_{t-r} = 0; x) = P(y_t < z_1|y_{t-1} < z_1, y_{t-2} < z_1, ..., y_{t-r+1} \leq z_1, y_{t-r} \leq z_0; x).
\]

■
Proof of Theorem 1

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

\[ E[|x_i\beta_i - x_i\beta_i'|^2] = E[|x_i(\beta_i - \beta_i')|^2] = (\beta_i - \beta_i')' E(x_i'x_i)(\beta_i - \beta_i') > 0, \]

where \( \beta_i \) is the true parameter vector and \( \beta_i,0 \) a parameter vector such that \( \beta_i,0 \neq \beta_i \).

Hence, \( x_i\beta_i \neq x_i\beta_i,0 \) with positive probability and since \( \Lambda(v) \) is strictly monotonic, we have \( \Lambda(x_i\beta_i) \neq \Lambda(x_i\beta_i,0) \) when \( x_i\beta_i \neq x_i\beta_i,0 \).

Condition (4) holds if \( E[|\log(f(S_i|S_{t-1} = 0... = S_{t-r} = 0; x_i; \beta_i)|] < \infty \) for all \( \beta_i \).

For the logistic function, it is easy to verify that

\[ |\ln \Lambda(v)| \leq |\ln \Lambda(0)| + |v|. \]

Furthermore, note that

\[
|\ln f(S_i|S_{t-1} = 0, ..., S_{t-r} = 0; x_i; \beta_i)| \leq |S_i| \ln (\Lambda(x_i\beta_i)| + |1 - S_i| \ln (1 - \Lambda(x_i\beta_i))|
\leq |\ln (\Lambda(x_i\beta_i))| + |\ln (1 - \Lambda(x_i\beta_i))|
\quad \text{(since } |S_i| \leq 1 \text{ and } |1 - S_i| \leq 1)\]
\leq 2[|\ln (\Lambda(0)) + ||x_i|| \times ||\beta_i||]
\quad \text{(due to the Cauchy-Schwartz inequality).}
\]

The nonsingularity of \( E(x_i'x_i') \) implies \( E(x_i^2_i) < \infty \) for all \( i \) and, therefore, \( E(||x_i||) < \infty \) and \( E(||x_i||) < \infty \), where \( ||.|| \) refers to the matrix norm. Thus, the nonsingularity of \( E(x_i'x_i') \) 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 \( B_i \) is taken to be \( \mathbb{R}^p \). Condition (2) is obviously satisfied. Regarding condition (3) note that since \( E[S_i|S_{t-1} = 0, ..., S_{t-r} = 0; x_i] = \Lambda(x_i\beta_i) \), we have \( E[s(w_i; \beta_i)|x_i] = 0 \) and, by the Law of Total Expectations \( E[s(w_i; \beta_i)] = 0 \).

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

\[
E\left[ \frac{\partial \ln L}{\partial \beta_i, \partial \beta_i'} \right] = -E\left[ \frac{\partial \ln L}{\partial \beta_i} \frac{\partial \ln L}{\partial \beta_i'} \right] + E\left[ \frac{1}{\ln L} \frac{\partial^2 \ln L}{\partial \beta_i \partial \beta_i'} \right],
\]

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:

\[-E[H(w_i; \beta_i)] = E[s(w_i; \beta_i) s(w_i; \beta_i)'],\]
where \(s(w_t; \beta_i)\) and \(H(w_t; \beta_i)\) are the functions defined in (6) and (7), respectively. Regarding the local dominance of the Hessian - condition (4) -, since \(\Lambda(x_t' \beta_i)[1 - \Lambda(x_t' \beta_i)] < 1\), we have \(\|H(w_t; \beta_i)\| \leq \|x_t x_t'\|\) for all \(\beta_i\). It can be shown that \(E[\|x_t x_t'\|] < \infty\) if \(E[x_t x_t']\) is nonsingular (and hence finite). Finally, condition (5) also requires that \(E[x_t x_t']\) is nonsingular.

**Proof of Theorem 4**

Under Assumption (A2), the joint stationarity of \(\{y_t, x_t\}\) implies the joint stationarity of \(\{S_t, x_t\}\), given the measurability of (1).

---

**Figure A.1**: Conditional estimated ET curves for \(z_0 = 1\), with overlapping block bootstrap 95% confidence intervals. For the baseline case the fiscal and monetary shocks equal their sample medians. An expansionary (contractionary) fiscal shock refers to the 0.8 (0.2) sample quantile and an expansionary (contractionary) monetary shock refers to the 0.2 (0.8) sample quantile. In Panel A the fiscal and monetary shocks equal their sample medians.
Table A.1: Bootstrap-based estimates for $p_1(x) := P(S_t = 0|S_{t-1} = 0; x)$.

Table: 1969q1-2008q4 ($T = 160$)

<table>
<thead>
<tr>
<th>$z_0$</th>
<th>$\hat{\beta}_{1,1}$</th>
<th>95% CI</th>
<th>$\hat{\beta}_{1,2}$</th>
<th>95% CI</th>
<th>$\hat{\beta}_{1,3}$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.742</td>
<td>2.669</td>
<td>2.823</td>
<td>-0.199</td>
<td>-0.240</td>
<td>-0.160</td>
</tr>
<tr>
<td>2</td>
<td>2.837</td>
<td>2.736</td>
<td>2.969</td>
<td>-0.196</td>
<td>-0.261</td>
<td>-0.133</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$z_0$</th>
<th>$\hat{\beta}_{1,1}$</th>
<th>95% CI</th>
<th>$\hat{\beta}_{1,2}$</th>
<th>95% CI</th>
<th>$\hat{\beta}_{1,3}$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.753</td>
<td>2.678</td>
<td>2.842</td>
<td>-0.210</td>
<td>-0.256</td>
<td>-0.170</td>
</tr>
<tr>
<td>2</td>
<td>2.838</td>
<td>2.735</td>
<td>2.975</td>
<td>-0.200</td>
<td>-0.267</td>
<td>-0.137</td>
</tr>
</tbody>
</table>

Note: The results in the upper panel of this Table are obtained using the updated Romer and Romer’s (2004) monetary policy shock series up to 2008q4 from Coibion et al. (2017), which we define as MS$_{RRt}$, and for the fiscal shock measure the Blanchard-Perotti shocks from Ramey and Zubairy (2018), defined as FS$_t$. The results in the lower panel of this Table use the same shocks as in the main text, MS$_t$, for the same sample range (1969q1 - 2008q4). In both cases, the fiscal and monetary shocks have the expected sign and the estimates are all statistically significant at the 5% significance level.
Table A.2: Bootstrap-based estimates for $p_1(x) := P(S_t = 0|S_{t-1} = 0; x)$ considering two fiscal shocks.

**Sample: 1969q1-2008q4 ($T=160$)**

<table>
<thead>
<tr>
<th>$z_0$</th>
<th>constant</th>
<th>$\hat{\beta}_{1,1}$</th>
<th>95% CI</th>
<th>$\hat{\beta}_{1,2}$</th>
<th>95% CI</th>
<th>$\hat{\beta}_{1,3}$</th>
<th>95% CI</th>
<th>$\hat{\beta}_{1,4}$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>2.755</td>
<td>2.682</td>
<td>2.841</td>
<td>-0.204</td>
<td>-0.249</td>
<td>-0.166</td>
<td>0.038</td>
<td>-0.003</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$z_0$</th>
<th>constant</th>
<th>$\hat{\beta}_{1,1}$</th>
<th>95% CI</th>
<th>$\hat{\beta}_{1,2}$</th>
<th>95% CI</th>
<th>$\hat{\beta}_{1,3}$</th>
<th>95% CI</th>
<th>$\hat{\beta}_{1,4}$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>2.768</td>
<td>2.690</td>
<td>2.856</td>
<td>-0.216</td>
<td>-0.264</td>
<td>-0.176</td>
<td>0.042</td>
<td>0.003</td>
</tr>
</tbody>
</table>

**Note:** The results in this Table are computed considering the narrative military (spending) news shocks ($FS_N_t$), in addition to the Blanchard-Perotti shocks ($FS_t$), both obtained from Ramey and Zubairy (2018). Since the military news shocks are not expected to be contemporaneously correlated with the Blanchard-Perotti shocks, we use these two fiscal shocks as regressors. Regarding the monetary shocks, since $MS_t$ and $MS_{RR_t}$ are correlated, we consider a specification with $MS_{RR_t}$ (upper panel) and another with $MS_t$ (bottom panel).

The $FS_N_t$ fiscal shock measure contains many zeros (83% of the sample observations), which causes some estimation difficulties. Since we are interested in transitional dynamics from $z_0$ to $z_1$, parameter estimation (for given $z_0$ and $z_1$) does not consider the entire sample (see (6) for details). Hence, it may happen that the number of cases with $S_t = 0$, $S_{t-1} = 0$ (see (1)) and nonzero $FS_N_t$ is not sufficient to properly estimate this fiscal shock coefficient. For instance, for $z_0 = 2$ there are only three cases in our sample that satisfy the conditions $S_t = S_{t-1} = 0$ and $FS_N_t \neq 0$. For this reason, we only present results for $z_0 = 1$. 