



BANCA D'ITALIA  
EUROSISTEMA

# Temi di Discussione

(Working Papers)

Bayesian analysis of coefficient instability in dynamic regressions

by Emanuela Ciapanna and Marco Taboga

November 2011

Number

836





BANCA D'ITALIA  
EUROSISTEMA

# Temi di discussione

(Working papers)

Bayesian analysis of coefficient instability in dynamic regressions

by Emanuela Ciapanna and Marco Taboga

Number 836 - November 2011

*The purpose of the Temi di discussione series is to promote the circulation of working papers prepared within the Bank of Italy or presented in Bank seminars by outside economists with the aim of stimulating comments and suggestions.*

*The views expressed in the articles are those of the authors and do not involve the responsibility of the Bank.*

*Editorial Board: SILVIA MAGRI, MASSIMO SBRACIA, LUISA CARPINELLI, EMANUELA CIAPANNA, ALESSANDRO NOTARPIETRO, PIETRO RIZZA, CONCETTA RONDINELLI, TIZIANO ROPELE, ANDREA SILVESTRINI, GIORDANO ZEVI.*

*Editorial Assistants: ROBERTO MARANO, NICOLETTA OLIVANTI.*

# BAYESIAN ANALYSIS OF COEFFICIENT INSTABILITY IN DYNAMIC REGRESSIONS

by Emanuela Ciapanna\* and Marco Taboga\*

## Abstract

This paper proposes a Bayesian regression model with time-varying coefficients (TVC) that makes it possible to estimate jointly the degree of instability and the time-path of regression coefficients. Thanks to its computational tractability, the model proves suitable to perform the first (to our knowledge) Monte Carlo study of the finite-sample properties of a TVC model. Under several specifications of the data generating process, the proposed model's estimation precision and forecasting accuracy compare favourably with those of other methods commonly used to deal with parameter instability. Furthermore, the TVC model leads to small losses of efficiency under the null of stability and it is robust to misspecification, providing a satisfactory performance also when regression coefficients experience discrete structural breaks. As a demonstrative application, we use our TVC model to estimate the exposures of S&P 500 stocks to market-wide risk factors: we find that a vast majority of stocks have time-varying risk exposures and that the TVC model helps to forecast these exposures more accurately.

**JEL Classification:** C11, C32, C50.

**Keywords:** time-varying regression, coefficient instability.

## Contents

1. Introduction.....	5
2. The Bayesian model .....	9
3. The specification of priors .....	16
4. Measures of (in)stability .....	21
5. Monte Carlo evidence.....	23
6. Empirical application: estimating common stocks' exposures to risk factors.....	34
7. Conclusions.....	35
8. Appendix.....	37
References .....	45
Tables .....	49

---

\* Bank of Italy, Economics, Research and International Relations.



# 1 Introduction<sup>1</sup>

There is widespread agreement that instability in regression coefficients represents a major challenge in empirical economics. In fact, many equilibrium relationships between economic variables are found to be unstable through time (e.g.: Stock and Watson - 1996).

There are two main approaches to address instability in regression coefficients:

1. formulating and estimating regression models under the hypothesis of constant coefficients, testing for the presence of structural breaks (e.g.: Chow - 1960, Brown, Durbin and Evans - 1975, Nyblom - 1989) and identifying the break-points (e.g.: Andrews, Lee, and Ploberger - 1996, Bai and Perron - 1998);
2. formulating regression models with time-varying coefficients (TVC) and estimating the path of their variation (e.g.: Doan, Litterman and Sims - 1984, Stock and Watson - 1996, Cogley and Sargent - 2001).

Approach (1) allows to search for time spans over which the hypothesis of constant coefficients is not rejected by the data. However, it can happen that regression coefficients change so frequently that the hypothesis of constant coefficients does not fit any time span (or only time spans that are too short to be of any interest to the econometrician). In these cases, approach (2) can be utilized, as it is suitable to deal also with frequently changing coefficients. On the other side of the coin, approach (2) often relies on dynamic specifications that are (at least in theory) not suitable to detect infrequent and abrupt changes in regression coefficients.

In the absence of strong priors about the ways in which relationships between variables change, the two approaches can arguably be considered complementary and it seems reasonable to use them in conjunction. However, approach (1) is apparently much more frequently utilized than approach (2) in empirical work (e.g.: Kapetanios - 2008).

---

<sup>1</sup>Any views expressed in this article are the authors' and do not necessarily represent those of the Bank of Italy. The bulk of this research project was done while Marco Taboga was visiting the Centre for Econometric Analysis at Cass Business School. We wish to thank for helpful comments Giovanni Urga and seminar participants at the Bank of Italy, at Cass Business School and at the Einaudi Institute for Economics and Finance, as well as participants at the Scottish Economic Society 2011 Conference.

One possible reason why TVC models are less popular is that tests for structural breaks are often quite easy to implement, while specifying and estimating TVC models is usually a difficult task that relies on complex and computationally intensive numerical techniques and requires careful specification of the dynamics of the coefficients. Even if the development of Markov chain Monte Carlo (MCMC) methods has somewhat facilitated the estimation of TVC models (e.g.: Carter and Kohn - 1994 and Chib and Greenberg - 1995), the technical skills and the computing time required by these techniques are still far superior to those required to estimate regressions with constant coefficients<sup>2</sup>.

In this paper, we propose a Bayesian TVC model that aims to fill this gap. The model has low computational requirements and allows to compute analytically the posterior probability that the regression is stable, the estimates of the regression coefficients and several other quantities of interest. Furthermore, it requires minimal input from the econometrician, in the sense that priors are specified automatically: in particular, the only inputs required from the econometrician are regressors and regressands, as in plain-vanilla OLS regressions with constant coefficients.

Another possible reason why TVC models are less popular than OLS-based alternatives is that the properties of the former are thus far largely unknown, while the latter have been extensively studied both theoretically (e.g.: Bai and Perron - 1998) and by means of Monte Carlo simulations (e.g.: Hansen - 2000 and Bai and Perron - 2006). Thanks to the computational tractability of our TVC model, we are able to perform the first (to our knowledge) Monte Carlo study of the finite sample properties of a TVC model.

The main goal of our Monte Carlo study is to address the concerns of an applied econometrician who suspects that the coefficients of a regression might be unstable, does not know what form of instability to expect and needs to decide what estimation strategy to adopt.

The first concern we address is loss of efficiency under the null of stability. *Suppose my data has indeed been generated by a regression with constant coefficients; how much do I lose, in terms of estimation precision and forecasting accuracy, when I estimate the regression using the TVC model in place of OLS?* Our results suggest that the

---

<sup>2</sup>In recent years, several empirical papers have successfully applied MCMC methods to the estimation of regression models with time varying parameters (e.g. Sargent, Williams and Zha - 2006, Canova and Ciccarelli - 2009, Canova and Gambetti - 2009, Koop, Leon-Gonzalez and Strachan - 2009, Cogley, Primiceri and Sargent - 2010). See also the survey by Koop and Korobilis (2009).



losses from using the TVC model are generally quite small and they are comparable to the losses from using frequentist breakpoint detection procedures, such as Bai and Perron's (1998 and 2003) sequential procedure and its model-averaging variant (Pesaran and Timmermann - 2007). Under most simulation scenarios, the mean squared estimation error increases by about 5 per cent when one of the proposed TVC estimators is used in place of OLS to estimate the coefficients of a stable regression.

Another concern is robustness to mis-specification. *Suppose my data has been generated by a regression with few discrete structural breaks; how much do I lose from using the TVC model instead of standard frequentist procedures for breakpoint detection?* Our Monte Carlo evidence indicates that also in this case the estimation precision and the forecasting accuracy of the TVC model are comparable to those of standard frequentist procedures.

Finally, a third concern is efficiency under the null of instability. *Even in the presence of frequently changing coefficients, does the TVC model provide better estimation precision and forecasting performance than other, possibly mis-specified, models?* We find that it generally does and that in some cases this gain in efficiency can be quite large (TVC can reduce the mean squared estimation error by up to 60 per cent with respect to the best performing OLS-based method).

All in all, the TVC model seems to be a valid complement to frequentist procedures for breakpoint detection, as the performances of the two approaches are, in general, comparable, but the TVC model fares better in the presence of frequently changing coefficients. There is, however, an important exception to this general result: when the regression includes a lag of the dependent variable and the autoregressive coefficient is near unity. In this case, the performance of the TVC model degrades steeply, and so, but to a lesser extent, does the performance of frequentist methods for breakpoint detection. We argue that this phenomenon is due to an identification problem (already pointed out in similar contexts by Hatanaka and Yamada - 1999 and Perron and Zhu - 2005) which can be alleviated by adding more regressors or increasing the sample size.

The Monte Carlo study is also complemented by a brief demonstration of how the TVC model can be applied to a real-world empirical problem. We consider a regression commonly employed to estimate how stock returns are related to market-wide risk factors. We find that the coefficients of this regression are unstable with high probability for a vast majority of the stocks included in the S&P 500 index. We

also find that the TVC model helps to better predict the exposures of these stocks to the risk factors.

Our model belongs to the family of Class I multi-process dynamic linear models defined by West and Harrison (1997). In our specification there is a single mixing parameter that takes on finitely many values between 0 and 1. The parameter measures the stability of regression coefficients: if it equals 0, then the regression is stable (coefficients are constant); the closer it is to 1, the more unstable coefficients are.

We propose two measures of stability that can be derived analytically from the posterior distribution of the mixing parameter, one based on credible intervals and one based on posterior odds ratios. We analyze the performance of a simple decision rule based on these measures of stability: "use OLS if they do not provide enough evidence of instability, otherwise use TVC". We find that such a decision rule performs well across different scenarios, leading to the smallest losses under the null of stability and still being able to produce satisfactory results when coefficients are indeed unstable.

Some features of our model are borrowed from existing TVC models (in particular Doan, Littermann and Sims - 1984, Stock and Watson - 1996, Cogley and Sargent - 2001), whereas other features are completely novel. First of all, we propose an extension of Zellner's (1986)  $g$ -prior to dynamic linear models. Thanks to this extension, posterior probabilities and coefficient estimates are invariant to re-scalings of the regressors<sup>3</sup>: this property is essential to obtain a completely automatic specification of priors. Another original feature of the model is the use of an invariant geometrically-spaced support for the prior distribution of the mixing parameter. We argue that this characteristic of the prior allows the model to capture both very low and very high degrees of coefficient instability, while retaining a considerable parsimony. Our modelling choices have two main practical consequences: 1) the priors are specified in a completely automatic way so that regressors and regressands are the only input required from the final user<sup>4</sup>; 2) the computational burden of the model is minimized,

---

<sup>3</sup>Before arriving to the specification of priors proposed in this paper we tried several other specifications and we found that results can indeed be quite sensitive to rescalings if one chooses other priors.

<sup>4</sup>A MATLAB function is made available on the internet at [http://www.statlect.com/time\\_varying\\_regression.htm](http://www.statlect.com/time_varying_regression.htm) The function can be called with the instruction:

$$\text{tvc}(y,X)$$

where  $y$  is a  $T \times 1$  vector of observations on the dependent variable and  $X$  is a  $T \times K$  matrix of regressors.

because analytical estimators are available both for the regression coefficients and for their degree of instability. To our knowledge, none of the existing models has these two characteristics, that allow to use the model in large-scale applications such as Monte Carlo simulations.

The paper is organized as follows: Section 2 presents the model; Section 3 describes the specification of priors; Section 4 introduces the two measures of (in)stability; Section 5 reports the results of the Monte Carlo experiments; Section 6 contains the empirical application; Section 7 concludes. Proofs and other technical details are relegated to the Appendix.

## 2 The Bayesian model

We consider a dynamic linear model (according to the definition given by West and Harrison - 1997) with time-varying regression coefficients:

$$y_t = x_t\beta_t + v_t \tag{1}$$

where  $x_t$  is a  $1 \times k$  vector of observable explanatory variables,  $\beta_t$  is a  $k \times 1$  vector of unobservable regression coefficients and  $v_t$  is an i.i.d. disturbance with normal distribution having zero mean and variance  $V$ . Time is indexed by  $t$  and goes from 1 to  $T$  ( $T$  is the last observation in the sample).

The vector of coefficients  $\beta_t$  is assumed to evolve according to the following equation:

$$\beta_t = \beta_{t-1} + w_t \tag{2}$$

where  $w_t$  is an i.i.d.  $k \times 1$  vector of disturbances having a multivariate normal distribution with zero mean and covariance matrix  $W$ .  $w_t$  is also contemporaneously and serially independent of  $v_t$ . The random walk hypothesis in (2), also adopted by Cogley and Sargent (2001) and Stock and Watson (1996), implies that changes in regression coefficients happen in an unpredictable fashion.

## 2.1 Notation

Let information available at time  $t$  be denoted by  $D_t$ .  $D_t$  is defined recursively by:

$$D_t = D_{t-1} \cup \{y_t, x_t\}$$

and  $D_0$  contains prior information on the parameters of the model (to be specified below).

We denote by  $(z | D_t)$  the distribution of a random vector  $z$ , given information at time  $t$  and by  $p(z | D_t)$  its probability density (or mass) function.

If a random vector  $z$  has a multivariate normal distribution with mean  $m$  and covariance matrix  $S$ , given  $D_t$ , we write:

$$(z | D_t) \sim N(m, S)$$

If a  $k \times 1$  random vector  $z$  has a multivariate Student's  $t$  distribution with mean  $m$ , scale matrix  $S$  and  $n$  degrees of freedom, we write:

$$(z | D_t) \sim T(m, S, n)$$

and its density is parametrized as follows:

$$p(z | D_t) \propto [n + (z - m)^\top S^{-1} (z - m)]^{\frac{k+n}{2}}$$

If  $z$  has a Gamma distribution with parameters  $V$  and  $n$ , we write

$$(z | D_t) \sim G(V, n)$$

and its density is parametrized as follows:

$$p(z | D_t) = \frac{(Vn/2)^{n/2} z^{n/2-1} \exp(-Vnz/2)}{\Gamma(n/2)}$$

Finally, define  $W^* = V^{-1}W$  and denote by  $X$  the design matrix

$$X = \left[ \begin{array}{ccc} x_1^\top & \dots & x_T^\top \end{array} \right]^\top$$

## 2.2 Structure of prior information and updating

In this subsection we state the main assumptions on the structure of prior information and we derive the formulae for updating the priors analytically.

The first set of assumptions regards  $\beta_1$ , the vector of regression coefficients at time  $t = 1$ , and  $V$ , the variance of the regression disturbances. We impose on  $\beta_1$  and  $V$  a conjugate normal/inverse-gamma prior<sup>5</sup>, i.e.:

- $\beta_1$  has a multivariate normal distribution conditional on  $V$ , with known mean  $\widehat{\beta}_{1,0}$  and covariance equal to  $V \cdot F_{\beta,1,0}$  where  $F_{\beta,1,0}$  is a known matrix;
- the reciprocal of  $V$  has a Gamma distribution, with known parameters  $\widehat{V}_0$  and  $n_0$ .

The second set of assumptions regards  $W$ , which is assumed to be proportional to the prior variance of  $\beta_1$ :

$$W = \lambda V F_{\beta,1,0} \quad (3)$$

where  $\lambda$  is a coefficient of proportionality<sup>6</sup>.

When  $\lambda = 0$ , the covariance matrix of  $w_t$  is zero and regression coefficients are stable. On the contrary, when  $\lambda > 0$ ,  $w_t$  has non-zero covariance matrix and the regression coefficients are unstable (i.e. they change through time). The higher  $\lambda$  is, the greater the variance of  $w_t$  is and the more unstable regression coefficients are.

The constant of proportionality  $\lambda$  is parametrized as:

$$\lambda = \lambda(\theta) \quad (4)$$

where  $\lambda(\theta)$  is a strictly increasing function and  $\theta$  is a random variable with finite support  $R_\theta$ :

$$R_\theta = \{\theta_1, \dots, \theta_q\} \subset [0, 1]$$

The prior probabilities of the  $q$  possible values of  $\theta$  are denoted by  $p_{0,1}, \dots, p_{0,q}$ . The discussion of how  $\theta_1, \dots, \theta_q$  and  $p_{0,1}, \dots, p_{0,q}$  are chosen is postponed to the

---

<sup>5</sup>This prior is frequently utilized in Bayesian regressions with constant coefficients (e.g. Hamilton - 1994)

<sup>6</sup>The assumption that  $W^* \propto F_{\beta,1,0}$  is made also by Doan, Littermann and Sims (1984) in their seminal paper on TVC models. However, in their model the coefficients  $\beta_t$  do not follow a random walk (they are mean reverting). They also use different priors: while we impose Zellner's  $g$ -prior on  $\beta_1$  (see section 3), they impose the Minnesota prior.

next section.

It is also assumed that  $\theta_1 = 0$  and  $\lambda(\theta_1) = 0$ . Therefore, regression coefficients are stable when  $\theta = \theta_1 = 0$  and unstable when  $\theta \neq \theta_1$  (the closer  $\theta$  is to 1, the more unstable regression coefficients are).

The assumptions on the priors and the initial information are summarized as follows:

**Assumption 1** *The priors on the unknown parameters are:*

$$\begin{aligned} (\beta_1 | D_0, V, \theta) &\sim N(\widehat{\beta}_{1,0}, VF_{\beta,1,0}) \\ (1/V | D_0, \theta) &\sim G(\widehat{V}_0, n_0) \\ p(\theta_i | D_0) &= p_{0,i}, i = 1, \dots, q \end{aligned}$$

and the initial information set is:

$$D_0 = \left\{ \widehat{\beta}_{1,0}, F_{\beta,1,0}, \widehat{V}_0, n_0, p_{0,1}, \dots, p_{0,q} \right\}$$

Given the above assumptions, the posterior distributions of the parameters of the regression can be calculated as follows:

**Proposition 2** *Let priors and initial information be as in Assumption 1. Let  $p_{t,i} = p(\theta = \theta_i | D_t)$ . Then:*

$$\begin{aligned} p(\beta_t | D_{t-1}) &= \sum_{i=1}^q p(\beta_t | \theta = \theta_i, D_{t-1}) p_{t-1,i} & (5) \\ p(y_t | D_{t-1}, x_t) &= \sum_{i=1}^q p(y_t | \theta = \theta_i, D_{t-1}, x_t) p_{t-1,i} \\ p(1/V | D_{t-1}) &= \sum_{i=1}^q p(1/V | \theta = \theta_i, D_{t-1}) p_{t-1,i} \\ p(\beta_t | D_t) &= \sum_{i=1}^q p(\beta_t | \theta = \theta_i, D_t) p_{t,i} \end{aligned}$$

where

$$\begin{aligned}
(\beta_t | \theta = \theta_i, D_{t-1}) &\sim T\left(\widehat{\beta}_{t,t-1,i}, \widehat{V}_{t-1,i} F_{\beta,t,t-1,i}, n_{t-1,i}\right) \\
(y_t | \theta = \theta_i, D_{t-1}, x_t) &\sim T\left(\widehat{y}_{t,t-1,i}, \widehat{V}_{t-1,i} F_{y,t,t-1,i}, n_{t-1,i}\right) \\
(1/V | \theta = \theta_i, D_{t-1}) &\sim G\left(\widehat{V}_{t-1,i}, n_{t-1,i}\right) \\
(\beta_t | \theta = \theta_i, D_t) &\sim T\left(\widehat{\beta}_{t,t,i}, \widehat{V}_{t,i} F_{\beta,t,t,i}, n_{t,i}\right)
\end{aligned} \tag{6}$$

and the parameters of the above distributions are obtained recursively as:

$$\begin{aligned}
\widehat{\beta}_{t,t-1,i} &= \widehat{\beta}_{t-1,t-1,i} & F_{\beta,t,t-1,i} &= F_{\beta,t-1,t-1,i} + \lambda(\theta_i) F_{\beta,1,0} \\
\widehat{y}_{t,t-1,i} &= x_t \widehat{\beta}_{t,t-1,i} & F_{y,t,t-1,i} &= 1 + x_t F_{\beta,t,t-1,i} x_t^\top \\
e_{t,i} &= y_t - \widehat{y}_{t,t-1,i} & P_{t,i} &= F_{\beta,t,t-1,i} x_t^\top / F_{y,t,t-1,i} \\
\widehat{\beta}_{t,t,i} &= \widehat{\beta}_{t,t-1,i} + P_{t,i} e_{t,i} & F_{\beta,t,t,i} &= F_{\beta,t,t-1,i} - P_{t,i} P_{t,i}^\top F_{y,t,t-1,i} \\
n_{t,i} &= n_{t-1,i} + 1 & \widehat{V}_{t,i} &= \frac{1}{n_{t,i}} \left( n_{t-1,i} \widehat{V}_{t-1,i} + \frac{e_{t,i}^2}{F_{y,t,t-1,i}} \right)
\end{aligned} \tag{7}$$

starting from the initial conditions  $\widehat{\beta}_{1,0,i} = \widehat{\beta}_{1,0}$ ,  $F_{\beta,1,0,i} = F_{\beta,1,0}$ ,  $\widehat{V}_{0,i} = \widehat{V}_0$  and  $n_{0,i} = n_0$ , while the mixing probabilities are obtained recursively as:

$$p_{t,i} = \frac{p_{t-1,i} p(y_t | \theta = \theta_i, D_{t-1}, x_t)}{\sum_{j=1}^q p_{t-1,j} p(y_t | \theta = \theta_j, D_{t-1}, x_t)} \tag{8}$$

starting from the prior probabilities  $p_{0,1}, \dots, p_{0,q}$ .

The updated mixing probabilities in the above proposition can be interpreted as posterior model probabilities, where a model is a TVC regression with fixed  $\theta$ . Hence, for example,  $p_{T,1}$  is the posterior probability of the regression model with stable coefficients ( $\theta = 0$ ). A crucial property of the framework we propose is that posterior model probabilities are known analytically: they can be computed exactly, without resorting to simulations.

In the above proposition, the priors on the regression coefficients  $\beta_t$  in a generic time period  $t$  are updated using only information received up to that same time  $t$ . However, after observing the whole sample (up to time  $T$ ), one might want to revise her priors on the regression coefficients  $\beta_t$  in previous time periods ( $t < T$ ), using the information subsequently received. This revision (usually referred to as smoothing) can be accomplished using the results of the following proposition:

**Proposition 3** *Let priors and initial information be as in Assumption 1. Then, for  $0 \leq \tau \leq T - 1$ :*

$$p(\beta_{T-\tau} | D_T) = \sum_{i=1}^q p(\beta_{T-\tau} | \theta = \theta_i, D_T) p_{T,i}$$

where

$$(\beta_{T-\tau} | \theta = \theta_i, D_T) \sim T\left(\widehat{\beta}_{T-\tau,T,i}, \widehat{V}_{T,i} F_{\beta,T-\tau,T,i}, n_{T,i}\right)$$

The mixing probabilities  $p_{T,i}$  and the parameters  $\widehat{V}_{T,i}$ ,  $n_{T,i}$  are obtained from the recursions in Proposition 2 while the parameters  $\widehat{\beta}_{T-\tau,T,i}$  and  $F_{\beta,T-\tau,T,i}$  are obtained from the following backward recursions:

$$\begin{aligned} \widehat{\beta}_{T-\tau,T,i} &= \widehat{\beta}_{T-\tau,T-\tau,i} + Q_{T-\tau,i} \left( \widehat{\beta}_{T-\tau+1,T,i} - \widehat{\beta}_{T-\tau+1,T-\tau,i} \right) \\ F_{\beta,T-\tau,T,i} &= F_{\beta,T-\tau,T-\tau,i} + Q_{T-\tau,i} (F_{\beta,T-\tau+1,T,i} - F_{\beta,T-\tau+1,T-\tau,i}) Q_{T-\tau,i}^\top \\ Q_{T-\tau,i} &= F_{\beta,T-\tau,T-\tau,i} (F_{\beta,T-\tau+1,T-\tau,i})^{-1} \end{aligned}$$

starting from  $\tau = 1$  and taking as final conditions the values  $\widehat{\beta}_{T-1,T-1,i}$ ,  $\widehat{\beta}_{T,T,i}$ ,  $\widehat{\beta}_{T,T-1,i}$ ,  $F_{\beta,T-1,T-1,i}$ ,  $F_{\beta,T,T,i}$  and  $F_{\beta,T,T-1,i}$  obtained from the recursions in Proposition 2.

Other important quantities of interest are known analytically, as shown by the following:



**Lemma 4** For  $1 \leq t \leq T$  and  $s \in \{t-1, t, T\}$ , the following equalities hold<sup>7</sup>:

$$\begin{aligned}
\mathbb{E}[\beta_t | D_s] &= \sum_{i=1}^q p_{s,i} \mathbb{E}[\beta_t | \theta = \theta_i, D_s] \\
\mathbb{E}[y_t | x_t, D_s] &= \sum_{i=1}^q p_{s,i} \mathbb{E}[y_t | x_t, \theta = \theta_i, D_s] \\
\text{Var}[\beta_t | D_s] &= \sum_{i=1}^q p_{s,i} \text{Var}[\beta_t | \theta = \theta_i, D_s] + \sum_{i=1}^q p_{s,i} \mathbb{E}[\beta_t | \theta = \theta_i, D_s] \mathbb{E}[\beta_t^\top | \theta = \theta_i, D_s] \\
&\quad - \left( \sum_{i=1}^q p_{s,i} \mathbb{E}[\beta_t | \theta = \theta_i, D_s] \right) \left( \sum_{i=1}^q p_{s,i} \mathbb{E}[\beta_t | \theta = \theta_i, D_s] \right)^\top \\
\text{Var}[y_t | x_t, D_s] &= \sum_{i=1}^q p_{s,i} \text{Var}[y_t | x_t, \theta = \theta_i, D_s] + \sum_{i=1}^q p_{s,i} \mathbb{E}[y_t | x_t, \theta = \theta_i, D_s]^2 \\
&\quad - \left( \sum_{i=1}^q p_{s,i} \mathbb{E}[y_t | x_t, \theta = \theta_i, D_s] \right)^2
\end{aligned}$$

where  $\mathbb{E}[\beta_t | \theta = \theta_i, D_s]$ ,  $\mathbb{E}[y_t | x_t, D_s]$ ,  $\text{Var}[\beta_t | D_s]$  and  $\text{Var}[y_t | x_t, D_s]$  can be calculated analytically for each  $\theta_i$  as in Propositions 2 and 3.

Thus, parameter estimates ( $\mathbb{E}[\beta_t | D_s]$ ) and predictions ( $\mathbb{E}[y_t | D_s]$ ) in any time period can be computed analytically and their variances are known in closed form. The probability distributions of  $\beta_t$  and  $y_t$  in a certain time period given information  $D_s$  are mixtures of Student's  $t$  distributions. Their quantiles are not known analytically, but they are easy to simulate by Monte Carlo methods. For example, if the distribution of  $\beta_T$  conditional on  $D_T$  is the object of interest, one can set up a Monte Carlo experiment where each simulation is conducted in two steps: 1) extract  $z$  from a uniform distribution on  $[0, 1]$ ; find  $k^*$  such that  $k^* = \arg \min \left\{ k : \sum_{i=1}^k p_{T,i} \geq z \right\}$ ; 2) given  $k^*$ , extract  $\beta_T$  from the Student's  $t$  distribution ( $\beta_T | \theta = \theta_{k^*}, D_T$ ), which is given by Proposition 2. The empirical distribution of the Monte Carlo simulations of  $\beta_T$  thus obtained is an estimate of the distribution of  $\beta_T$  conditional on  $D_T$ .

---

<sup>7</sup>These are trivial consequences of the linearity of the integral (for the expected values) and of the law of total variance (for the variances).

### 3 The specification of priors

Our specification of priors aims to be:

1. *objective*, in the sense that it does not require elicitation of subjective priors;
2. *fully automatic*, in the sense that the model necessitates no inputs from the econometrician other than regressors and regressands, as in plain-vanilla OLS regressions with constant coefficients.

The above goals are pursued by extending Zellner's (1986)  $g$ -prior to TVC models and by parametrizing  $\lambda(\theta)$  in such a way that the support of  $\theta$  is invariant (it needs not be specified on a case-by-case basis).

#### 3.1 The prior mean and variance of the coefficients

We use a version of Zellner's (1986)  $g$ -prior for the prior distribution of the regression coefficients at time  $t = 1$ :

**Assumption 5** *The prior mean is zero, corresponding to a prior belief of no predictability:*

$$\widehat{\beta}_{1,0} = 0 \tag{9}$$

*while the prior covariance matrix is proportional to  $(X^\top X)^{-1}$ :*

$$F_{\beta,1,0} = g (X^\top X)^{-1} \tag{10}$$

*where  $g$  is a coefficient of proportionality.*

Zellner's (1986)  $g$ -prior is widely used in model selection and model averaging problems similar to ours (we have a range of regression models featuring different degrees of instability), because it greatly reduces the sensitivity of posterior model probabilities to the specification of prior distributions (Fernandez, Ley and Steel - 2001), thus helping to keep the analysis as objective as possible. Furthermore, Zellner's (1986)  $g$ -prior has a straightforward interpretation: it can be interpreted as information provided by a conceptual sample having the same design matrix  $X$  as the current sample (Zellner 1986; George and McCulloch 1997; Smith and Kohn 1996).

To keep the prior relatively uninformative, we follow Kass and Wasserman (1995) and choose  $g = T$  (see also Shively, Kohn and Wood - 1999):

**Assumption 6** *The coefficient of proportionality is  $g = T$ .*

Thus, the amount of prior information (in the Fisher sense) about the coefficients is equal to the amount of average information contained in one observation from the sample.

**Remark 7** *Given that  $W^* = \lambda(\theta) F_{\beta,1,0}$  (equations 3 and 4), Zellner's prior (10) implies that also the covariance matrix of  $w_t$  is proportional to  $(X^\top X)^{-1}$ :*

$$W^* \propto (X^\top X)^{-1}$$

This proportionality condition has been imposed in a TVC model also by Stock and Watson<sup>8</sup> (1996), who borrow it from Nyblom (1989). A similar hypothesis is adopted also by Cogley and Sargent<sup>9</sup> (2001).

**Remark 8** *Given (10) and (9), all the coefficients  $\beta_t$  have zero prior mean and covariance proportional to  $(X^\top X)^{-1}$  conditional on  $D_0$ :*

$$\begin{aligned} \mathbb{E}[\beta_t | D_0, V, \theta] &= 0, \quad t = 1, \dots, T, \quad \forall \theta \\ \text{Var}[\beta_t | D_0, V, \theta] &= [T + (t - 1)\lambda(\theta)] V (X^\top X)^{-1}, \quad t = 1, \dots, T, \quad \forall \theta \end{aligned}$$

This property will be used later, together with other properties of the priors, to prove that posterior model probabilities are scale invariant in the covariates.

### 3.2 The variance parameters $\widehat{V}_0$ and $n_0$

In objective Bayesian analyses, the prior usually assigned to  $V$  in conjunction with Zellner's (1986)  $g$ -prior (e.g.: Liang et al. - 2008) is the improper prior:

$$p(V | D_0, \theta) \propto V^{-1}$$

---

<sup>8</sup>However, they assume that  $F_{\beta,1,0}$  is proportional to the identity matrix, while we assume that also  $F_{\beta,1,0}$  is proportional to  $(X^\top X)^{-1}$ . Furthermore, they do not estimate  $V$ . Their analysis is focused on the one-step-ahead predictions of  $y_t$ , which can be computed without knowing  $V$ . They approach the estimation of  $\lambda$  in a number of different ways, but none of them allows to derive analytically a posterior distribution for  $\lambda$ .

<sup>9</sup>In their model the prior covariance of  $w_t$  is proportional to  $(X^\top X)^{-1}$ , but  $X$  is the design matrix of a pre-sample not used for the estimation of the model.

With this choice, the updating equations in Proposition 2 would have to be replaced with a different set of updating equations until reaching the first non-zero observation of  $y_t$  (see e.g. West and Harrison - 1997). Furthermore, the updating of posterior probabilities would be slightly more complicated. To avoid the subtleties involved in using an improper prior, we adopt a simpler procedure, which yields almost identical results in reasonably sized samples:

**Assumption 9** *The first observation in the sample (denote it by  $y_0$ ) is used to form the prior on  $V$  :*

$$\begin{aligned}\widehat{V}_0 &= y_0^2 \\ n_0 &= 1\end{aligned}$$

After using it to form the prior, we discard the first observation and start updating the equations (5)-(8) from the following observation. If the first observation is zero ( $y_0 = 0$ ) we discard it and use the next to form the prior (or repeat until we find the next non-zero observation).

### 3.3 The mixing parameter $\theta$

We have assumed that  $W^* = \lambda(\theta) F_{\beta,1,0}$  where  $\theta$  is a random variable having finite support  $R_\theta = \{\theta_1, \dots, \theta_q\} \subset [0, 1]$ ,  $\theta_1 = 0$  and  $\lambda(\theta)$  is strictly increasing in  $\theta$  and such that  $\lambda(\theta_1) = 0$ . We now propose a specification of the function  $\lambda(\theta)$  that satisfies the above requirements and allows for an intuitive interpretation of the parameter  $\theta$ , while also facilitating the specification of a prior distribution for  $\theta$ .

First, note that:

$$y_t = x_t \beta_{t-1} + x_t w_t + v_t$$

Hence, given  $\lambda$ ,  $x_t$  and the initial information  $D_0$ , the variance generated by innovations at time  $t$  is:

$$\text{Var} [x_t w_t + v_t | x_t, D_0, \lambda] = \lambda \widehat{V}_0 x_t F_{\beta,1,0} x_t^\top + \widehat{V}_0$$

**Assumption 10**  *$\theta$  is the fraction of  $\text{Var} [x_t w_t + v_t | x_t, D_0, \lambda]$  generated on average*

by innovations to the regression coefficients:

$$\theta = \frac{\frac{1}{T} \sum_{t=1}^T \text{Var} [x_t w_t | x_t, D_0, \lambda]}{\frac{1}{T} \sum_{t=1}^T \text{Var} [x_t w_t + v_t | x_t, D_0, \lambda]}$$

Given this assumption on  $\theta$ , it is immediate to prove that

$$\lambda(\theta) = \frac{\theta}{\omega - \theta\omega}$$

where

$$\omega = \frac{1}{T} \sum_{t=1}^T x_t F_{\beta,1,0} x_t^\top$$

$\lambda(\theta)$  is strictly increasing in  $\theta$  and such that  $\lambda(0) = 0$ , as required. Hence, when  $\theta = 0$  the regression has stable coefficients. Furthermore, by an appropriate choice of  $\theta$ , any degree of coefficient instability can be reproduced (when  $\theta$  tends to 1,  $\lambda$  approaches infinity).

As far as the support of  $\theta$  is concerned, we make the following assumption:

**Assumption 11** *The support of  $\theta$  is a geometrically spaced grid, consisting of  $q$  points:*

$$R_\theta = \{0, \theta_{\max} c^{q-2}, \theta_{\max} c^{q-1}, \dots, \theta_{\max} c^1, \theta_{\max}\}$$

where  $0 \leq \theta_{\max} < 1$  and  $0 < c < 1$ .

Notice that  $\theta_{\max}$  cannot be chosen to be exactly equal to 1 (because  $\lambda(1) = \infty$ ), but it can be set equal to any number arbitrarily close to 1.

Using a geometrically spaced grid is the natural choice when the order of magnitude of a parameter is unknown (e.g.: Guerre and Lavergne - 2005, Horowitz and Spokoiny - 2001, Lepski, Mammen and Spokoiny - 1997 and Spokoiny - 2001): in our model, it allows to simultaneously consider both regressions that are very close to being stable and regressions that are far from being stable, without requiring too fine a grid<sup>10</sup>.

If the geometric grid is considered as an approximation of a finer set of points (possibly a continuum), the geometric spacing ensures that the maximum relative

---

<sup>10</sup>For example, in the empirical part of the paper, setting  $q = 100$  and  $c = 0.9$ , we are able to simultaneously consider 5 different orders of magnitude of instability. With the same number of points  $q$  and an arithmetic grid, we would have been able to consider only 2 orders.

round-off error is constant on all subintervals  $[\theta_i, \theta_{i+1}]$  such that  $1 < i < q$ . The maximum relative round-off error is approximately  $\frac{1-c}{2}$  on these subintervals and it can be controlled by an appropriate choice of  $c$ . On the contrary, the maximum relative round-off error cannot be controlled (it always equals 1) on the subinterval  $[\theta_0, \theta_1]$ , because the latter contains the point  $\theta_0 = 0$ . Only the absolute round-off error (equal to  $\frac{\theta_{\max}}{2}c^{q-2}$ ) can be controlled on  $[\theta_0, \theta_1]$ , by an appropriate choice of  $q$ . Therefore, setting the two parameters  $c$  and  $q$  can be assimilated to setting the absolute and relative error tolerance in a numerical approximation problem.

Assuming prior ignorance on the order of magnitude of  $\theta$ , we assign equal probability to each point in the grid:

**Assumption 12** *The prior mixing probabilities are assumed to be:*

$$p(\theta_i | D_0) = q^{-1}, i = 1, \dots, q$$

Note that, given the above choices, the prior on  $\theta$  and its support are invariant, in the sense that they do not depend on any specific characteristic of the data to be analyzed, but they depend only on the maximum percentage round-off error  $\frac{1-c}{2}$ . As a consequence, they allow the specification of priors to remain fully automatic.

### 3.4 Scale invariance

A crucial property of the automatic specification of priors proposed in the previous sections is that it guarantees scale invariance. The scale invariance property is satisfied if multiplying the regressors by an invertible matrix  $R$ , the posterior distribution of the coefficients is re-scaled accordingly<sup>11</sup> (it is multiplied by  $R^{-1}$ ). Virtually all the TVC models we have found in the literature do not satisfy the scale invariance property, in the sense that they do not contemplate a mechanism to guarantee scale invariance by automatically re-scaling priors when the scale of regressors is changed. Although scale invariance might seem a trivial property, it is indispensable to achieve one of the main goals of this paper: having a completely automatic model that requires only regressors and regressands as inputs from the econometrician. Furthermore, it guarantees replicability of results: two researchers using the same data, but on different scales, will obtain the same results.

---

<sup>11</sup>For example, when a regressor is multiplied by 100, the posterior mean of the coefficient multiplying that regressor is divided by 100.

Scale invariance is formally defined as follows:

**Definition 13** *Given the initial information set  $D_0$ , the information sets*

$$D_t = D_{t-1} \cup \{y_t, x_t\} \quad t = 1, \dots, T$$

*and a full-rank  $k \times k$  matrix  $R$ , an initial information set  $D_0^*$  is said to be  $R$ -scale invariant with respect to  $D_0$  if and only if:*

$$(\beta_t | D_s^*) = (R^{-1}\beta_t | D_s) \quad \forall s, t \leq T$$

where

$$D_t^* = D_{t-1}^* \cup \{y_t, x_t R\} \quad t = 1, \dots, T$$

Note that the initial information set  $D_0$ , which contains the priors, is automatically specified as a function of  $y_0$  and  $(X^\top X)^{-1}$ . We can write:

$$D_0 = D \left( y_0, (X^\top X)^{-1} \right) \quad (11)$$

The following proposition, proved in the Appendix, shows in what sense our TVC model is scale-invariant:

**Proposition 14** *For any full-rank  $k \times k$  matrix  $R$ , the initial information set  $D_0^*$  defined by:*

$$D_0^* = D \left( y_0, (R^\top X^\top X R)^{-1} \right)$$

*is  $R$ -scale invariant with respect to the initial information set  $D_0$ , as defined in (11).*

## 4 Measures of (in)stability

After computing the posterior distribution of  $\theta$ , a researcher might naturally ask: *how much evidence did the data provide against the hypothesis of stability?* Here, we discuss some possible ways to answer this question.

The crudest way to evaluate instability is to look at the posterior probability that  $\theta = 0$ . The closer to 1 this probability is, the more evidence of stability we have. However, a low posterior probability that  $\theta = 0$  does not necessarily constitute overwhelming evidence of instability. It might simply be the case that the sample

is not large enough to satisfactorily discriminate, *a posteriori*, between stable and unstable regressions: in such cases, even if the true regression is stable, unstable regressions might be assigned posterior probabilities that are only marginally lower than the probability of the stable one. Furthermore, if  $R_\theta$  contains a great number of points, it can happen that the posterior probability that  $\theta = 0$  is close to zero, but still much higher than the posterior probability of all the other points.

We propose two measures of stability to help circumvent the above shortcomings.

The first measure of stability, denoted by  $\Pi$ , is based on credible intervals (e.g.: Robert - 2007):

**Definition 15 ( $\Pi$ -stability)** *Let  $H_\theta$  be a higher posterior probability set defined as follows:*<sup>12</sup>

$$H_\theta = \{\theta_i \in R_\theta : p(\theta = \theta_i | D_T) > p(\theta = 0 | D_T)\}$$

*The stability measure  $\Pi$  is defined by:*

$$\Pi = 1 - \frac{\sum_{\theta_i \in H_\theta} p(\theta = \theta_i | D_T)}{\sum_{\theta_i \neq 0} p(\theta = \theta_i | D_T)},$$

*where we adopt the convention  $0/0 = 0$ .*

When  $\Pi = 1$ ,  $\theta = 0$  is a mode of the posterior distribution of  $\theta$ : we attach to the hypothesis of stability a posterior probability that is at least as high as the posterior probability of any alternative hypothesis of instability. On the contrary, when  $\Pi = 0$ , the posterior probability assigned to the hypothesis of stability is so low that all unstable models are more likely than the stable one, *a posteriori*. In the intermediate cases ( $0 < \Pi < 1$ ),  $\Pi$  provides a measure of how far the hypothesis of stability is from being the most likely hypothesis (the lower  $\Pi$ , the less likely stability is).

The second measure of stability, denoted by  $\pi$ , is constructed as a posterior odds ratio and it is based on the probability of the posterior mode of  $\theta$ .

**Definition 16 ( $\pi$ -stability)** *Let  $p^*$  be the probability of (one of) the mode(s) of the posterior distribution of  $\theta$ :*

$$p^* = \max(\{p(\theta = \theta_i | D_T) : \theta_i \in R_\theta\})$$

---

<sup>12</sup> $H_\theta$  contains all points of  $R_\theta$  having higher posterior probability than  $\theta = 0$  (recall that  $\theta = 0$  means that regression coefficients are stable).



The stability measure  $\pi$  is defined by:

$$\pi = \frac{p(\theta = 0 | D_T)}{p^*}$$

As with the previously proposed measure, when  $\pi = 1$ ,  $\theta = 0$  is a mode of the posterior distribution of  $\theta$  and stability is the most likely hypothesis, *a posteriori*. On the contrary, the closer  $\pi$  is to zero, the less likely stability is, when compared with the most likely hypothesis. For example, when  $\pi = 1/10$ , there is an unstable regression that is 10 times more likely than the stable one.

Both measures of stability ( $\Pi$  and  $\pi$ ) can be used to make decisions. For example, one can fix a threshold  $\gamma$  and decide to reject the hypothesis of stability if the measure of stability is below the threshold ( $\Pi < \gamma$  or  $\pi < \gamma$ ). In case  $\Pi$  is used, the procedure can be assimilated to a frequentist test of hypothesis, where  $1 - \gamma$  represents the level of confidence.  $\Pi$  can be interpreted as a sort of Bayesian  $p$ -value (e.g.: Robert - 2007): the lower  $\Pi$  is, the higher is the confidence with which we can reject the hypothesis of stability<sup>13</sup>. In case  $\pi$  is used, one can resort to Jeffreys' (1961) scale to qualitatively assess the strength of the evidence against the hypothesis of stability (e.g.: substantial evidence if  $\frac{1}{3} \leq \pi < \frac{1}{10}$ , strong evidence if  $\frac{1}{10} \leq \pi < \frac{1}{30}$ , very strong evidence if  $\frac{1}{30} \leq \pi < \frac{1}{100}$ ).

In the next section we explore the consequences of using these decision rules to decide whether to estimate a regression by OLS or by TVC.

## 5 Monte Carlo evidence

### 5.1 Performance when the DGP is a stable regression

In this subsection we present the results of a set of Monte Carlo simulations aimed at evaluating how much efficiency is lost when a stable regression is estimated with our TVC model. We compare the forecasting performance and the estimation precision of the TVC model with those of plain vanilla OLS and of a standard frequentist procedure used to identify breakpoints and estimate regression coefficients in the

---

<sup>13</sup>Note, however, that the parallelism can be misleading, as Bayesian  $p$ -values have frequentist validity only in special cases. Ghosh and Mukerjee (1993), Mukerjee and Dey (1993), Datta and Ghosh (1995) and Datta (1996) provide conditions that priors have to satisfy in order for Bayesian  $p$ -values to have also frequentist validity.

presence of structural breaks. In particular, we consider the performance of Bai and Perron's (1998 and 2003) sequential procedure, as implemented by Pesaran and Timmermann (2002 and 2007).

For our Monte Carlo experiments, we adapt a design that has already been employed in the literature on parameter instability (Hansen - 2000).

The design is as follows:

- *Data generating process:*  $y_t$  is generated according to:

$$y_t = \rho y_{t-1} + u_{t-1} + v_t$$

where  $y_0 = 0$ ,  $u_t \sim T(0, 1, 5)$  i.i.d.,  $v_t \sim N(0, 1)$  i.i.d. and  $u_t$  and  $z_t$  are serially and cross-sectionally independent.

- *Estimated equations:* two equations are estimated. In the first case, a constant and the first lags of  $y_t$  and  $u_t$  are included in the set of regressors; hence, the estimated model is (1), where

$$x_t = \begin{bmatrix} 1 & y_{t-1} & u_{t-1} \end{bmatrix}$$

In the second case, a constant and the first three lags of  $y_t$  and  $u_t$  are included in the set of regressors; hence, the estimated model is (1), where

$$x_t = \begin{bmatrix} 1 & y_{t-1} & y_{t-2} & y_{t-3} & u_{t-1} & u_{t-2} & u_{t-3} \end{bmatrix}$$

- *Parameters of the design:* simulations are conducted for three different sample sizes ( $T = 100, 200, 500$ ), four different values of the autoregressive coefficient ( $\rho = 0, 0.50, 0.80, 0.99$ ) and the two estimated equations detailed above, for a total of 24 experiments.

Each Monte Carlo experiment consists of 10,000 simulations.

The loss in estimation precision is evaluated comparing the estimate of the coefficient vector at time  $T$  (denote it by  $\tilde{\beta}_T$ ) with its true value. We consider seven different estimates:

- model averaging (TVC-MA) estimates, where:

$$\tilde{\beta}_T = \mathbb{E}[\beta_T | D_T] = \sum_{i=1}^q p_{T,i} \mathbb{E}[\beta_T | D_T, \theta = \theta_i]$$

- model selection (TVC-MS) estimates, where:

$$\tilde{\beta}_T = \mathbb{E}[\beta_T | D_T, \theta = \theta_{j^*}]$$

and

$$j^* = \arg \max_j p_{T,j}$$

i.e. only the model with the highest posterior probability is used to make predictions;

- estimates obtained from the regression model with stable coefficients when  $\Pi \geq 0.1$  and from model averaging when  $\Pi < 0.1$  (denoted by TVC-II):

$$\tilde{\beta}_T = \begin{cases} \mathbb{E}[\beta_T | D_T, \theta = 0] & \text{if } \Pi \geq 0.1 \\ \mathbb{E}[\beta_T | D_T] & \text{if } \Pi < 0.1 \end{cases}$$

i.e. coefficients are estimated with the TVC model only if there is enough evidence of instability ( $\Pi < 0.1$ ); otherwise, the standard OLS regression is used. This is intended to reproduce the outcomes of a decision rule whereby the econometrician uses the TVC model only if the TVC model itself provides enough evidence that OLS is inadequate;

- estimates obtained from the regression model with stable coefficients when  $\pi \geq 0.1$  and from model averaging when  $\pi < 0.1$  (denoted by TVC- $\pi$ ):

$$\tilde{\beta}_T = \begin{cases} \mathbb{E}[\beta_T | D_T, \theta = 0] & \text{if } \pi \geq 0.1 \\ \mathbb{E}[\beta_T | D_T] & \text{if } \pi < 0.1 \end{cases}$$

This estimator is similar to the previous one, but  $\pi$  is used in place of  $\Pi$  to decide whether there is enough evidence of instability;

- estimates obtained from the regression model with stable coefficients (OLS):

$$\tilde{\beta}_T = \mathbb{E}[\beta_T | D_T, \theta = 0]$$

- OLS estimates obtained from Bai and Perron's (1998 and 2003) sequential<sup>14</sup> procedure (denoted by BP), using the SIC criterion to choose the number of breakpoints (Pesaran and Timmermann - 2002 and 2007). If  $\tilde{\tau}$  is the last estimated breakpoint date in the sample, then  $\tilde{\beta}_T$  is the OLS estimate of  $\beta_T$  obtained using all the sample points from  $\tilde{\tau}$  to  $T$ ;
- estimates obtained from Pesaran and Timmermann's (2007) model-averaging procedure (denoted by BP-MA): the location of the last breakpoint is estimated with Bai and Perron's procedure (as in the point above); if  $\tilde{\tau}$  is the last estimated breakpoint date in the sample, then:

$$\tilde{\beta}_T = \sum_{\tau=1}^{\tilde{\tau}} w_{\tau} \tilde{\beta}_{T,\tau}$$

where  $\tilde{\beta}_{T,\tau}$  is the OLS estimate of  $\beta_T$  obtained using all the sample points from  $\tau$  to  $T$ ;  $w_{\tau}$  is a weight proportional to the inverse of the mean squared prediction error committed when using only the sample points from  $\tau$  onwards to estimate the regression and predict  $y_t$  ( $\tau + k + 1 \leq t \leq T$ ).

The Monte Carlo replications are used to estimate the mean squared error of the coefficient estimates:

$$MSE_j^{\beta} = \mathbb{E} \left[ \left\| \beta_T - \tilde{\beta}_T \right\|^2 \right]$$

where  $\| \cdot \|$  is the Euclidean norm and  $j = \text{TVC-MA, TVC-MS, TVC-II, TVC-}\pi, \text{OLS, BP, BP-MA}$  depending on which of the above methods has been used to estimate  $\beta_T$ .

The two parameters regulating the granularity of the grid for  $\theta$  are chosen as follows:  $q = 100$  and  $c = 0.9$ . To avoid degeneracies, rather than setting  $\theta_{\max} = 1$

---

<sup>14</sup>We estimate the breakpoint dates sequentially rather than simultaneously to achieve a reasonable computational speed in our Monte Carlo simulations. Denote by  $\nu_S$  the number of breakpoints estimated by the sequential procedure and by  $\nu_{\sigma}$  the number estimated by the simultaneous procedure. Given that we are using the SIC criterion to choose the number of points, if  $\nu_{\sigma} \leq 1$ , then  $\nu_S = \nu_{\sigma}$ ; otherwise, if  $\nu_{\sigma} > 1$ , then  $\nu_S \leq \nu_{\sigma}$ . Therefore, in our Montecarlo simulations (where the true number of breakpoints is either 0 or 1), the sequential procedure provides a better estimate of the number of breakpoints than the simultaneous procedure.

(the theoretical upper bound on  $\theta$ ), we choose a value that is numerically close to 1 ( $\theta_{\max} = 0.999$ ). Thus, the relative round-off error is bounded at 5 per cent and the model is able to detect degrees of instability as low as  $\theta \simeq 3 \cdot 10^{-5}$  (for concreteness, this means that coefficient instability can be detected by the model also in cases in which less than 0.01 per cent of total innovation variance is generated by coefficient instability).

Panel A of Table 1 reports the Monte Carlo estimates of  $MSE_j^\beta$  for the case in which  $x_t$  includes only the first lags of  $y_t$  and  $u_t$ . Not surprisingly, the smallest MSE is in all cases achieved by the OLS estimates. As anticipated in the introduction, there are significant differences between the case in which the autoregressive component is very persistent ( $\rho = 0.99$ ) and the other cases ( $\rho = 0, 0.50, 0.80$ ). In the latter cases, the TVC- $\pi$  coefficient estimates are those that yield the smallest increase in MSE with respect to OLS (in most cases under 5 per cent). The performance of BP-MA is the second best, being only slightly inferior to that of TVC- $\pi$ , but slightly superior to that of TVC-II.  $MSE_{TVC-MA}^\beta$  and  $MSE_{TVC-MS}^\beta$  are roughly between 20 and 60 per cent higher than  $MSE_{TVC-OLS}^\beta$ , while  $MSE_{TVC-BP}^\beta$  is on average equal to several multiples of  $MSE_{TVC-OLS}^\beta$ . Qualitatively speaking, the loss in precision from using TVC-II, TVC- $\pi$  and BP-MA is almost negligible, while there is a severe loss using BP and a moderate loss using TVC-MA and TVC-MS. In the case in which  $\rho = 0.99$ , results are very different: on average,  $MSE_{TVC}^\beta$  (all four kinds of TVC) and  $MSE_{BP}^\beta$  become almost two orders of magnitude greater than  $MSE_{OLS}^\beta$ , while  $MSE_{BP-MA}^\beta$  remains comparable to  $MSE_{OLS}^\beta$  (although there is a worsening with respect to the case of low persistence).

The unsatisfactory performance of the TVC and BP estimates in the case of high persistence can arguably be explained by an identification problem. In the unit root case, the regression generating the data is:

$$y_t = y_{t-1} + u_{t-1} + v_t$$

For any  $\alpha < 1$ , it can be rewritten as:

$$y_t = \mu_t + \alpha y_{t-1} + u_{t-1} + v_t$$

where  $\mu_t = (1 - \alpha) y_{t-1}$  is an intercept following a random walk. Furthermore, its innovations ( $\mu_t - \mu_{t-1}$ ) are contemporaneously independent of the innovations  $v_t$ . There-

fore, if the estimated equation includes a constant and time-varying coefficients are not ruled out, it is not possible to identify whether the regression has a unit root and stable coefficients or has a stationary autoregressive component and a time-varying intercept<sup>15</sup>. When  $\rho$  is near unity, identification is possible, but it will presumably be weak, giving rise to very imprecise estimates of the coefficients and of their degree of stability. Note that the two equivalent (and unidentified) representations above obviously yield the same one-step-ahead forecasts of  $y_t$ . Therefore, if our conjecture that this weak identification problem is affecting our results is correct, we should find that the out-of-sample forecasts of  $y_t$  produced by the TVC model are not as unsatisfactory as its coefficient estimates. This is exactly what we find and document in the last part of this subsection.

Panel B of Table 1 reports the Monte Carlo estimates of  $MSE_j^\beta$  for the case in which  $x_t$  includes three lags of  $y_t$  and  $u_t$ . In the case of low persistence, the BP-MA estimates are those that achieve the smallest increase in MSE with respect to the OLS estimates (on average below 2 per cent). The performance of the TVC- $\pi$  estimates is only slightly inferior (around 3 percent increase in MSE with respect to OLS). All the other estimates (TVC-MA, TVC-MS, TVC-II and BP) are somewhat less efficient, but their MSEs seldom exceed those of the OLS estimates by more than 30%. As far as the highly persistent case ( $\rho = 0.99$ ) is concerned, we again observe a degradation in the performance of the TVC and (to a lesser extent) of the BP estimates. However, the degradation is less severe than the one observed in the case of fewer regressors. Intuitively, adding more regressors (even if their coefficients are 0) helps to alleviate the identification problem discussed before, because the added regressors have stable coefficients and hence help to pin down the stable representation of the regression.

The loss in forecasting performance is evaluated using a single out-of-sample prediction for each replication. In each replication,  $T + 1$  observations are generated, the first  $T$  are used to update the priors, the vector of regressors  $x_{T+1}$  is used to predict  $y_{T+1}$  and the prediction (denote it by  $\tilde{y}_{T+1}$ ) is compared to the actual value  $y_{T+1}$ . As for coefficient estimates, we consider seven different predictions:

---

<sup>15</sup>This identification problem is discussed in a very similar context by Hatanaka and Yamada (1999) and Perron and Zhu (2005).

- model averaging (TVC-MA) predictions, where:

$$\tilde{y}_{T+1} = \text{E} [y_{T+1} | D_T, x_{T+1}] = \sum_{i=1}^q p_{T,i} \text{E} [y_{T+1} | D_T, x_{T+1}, \theta = \theta_i]$$

- model selection (TVC-MS) predictions, where:

$$\tilde{y}_{T+1} = \text{E} [y_{T+1} | D_T, x_{T+1}, \theta = \theta_{j^*}]$$

- predictions generated by the regression model with stable coefficients when  $\Pi \geq 0.1$  and by model averaging when  $\Pi < 0.1$  (denoted by TVC- $\Pi$ ):

$$\tilde{y}_{T+1} = \begin{cases} \text{E} [y_{T+1} | D_T, x_{T+1}, \theta = 0] & \text{if } \Pi \geq 0.1 \\ \text{E} [y_{T+1} | D_T, x_{T+1}] & \text{if } \Pi < 0.1 \end{cases}$$

- predictions generated by the regression model with stable coefficients when  $\pi \geq 0.1$  and by model averaging when  $\pi < 0.1$  (denoted by TVC- $\pi$ ):

$$\tilde{y}_{T+1} = \begin{cases} \text{E} [y_{T+1} | D_T, x_{T+1}, \theta = 0] & \text{if } \pi \geq 0.1 \\ \text{E} [y_{T+1} | D_T, x_{T+1}] & \text{if } \pi < 0.1 \end{cases}$$

- predictions generated by the regression model with stable coefficients (OLS):

$$\tilde{y}_{T+1} = \text{E} [y_{T+1} | D_T, x_{T+1}, \theta = 0]$$

- predictions obtained from Bai and Perron's sequential procedure (BP); if  $\tilde{\beta}_T$  is the BP estimate of  $\beta_T$  (see above), then:

$$\tilde{y}_{T+1} = x_{T+1} \tilde{\beta}_T$$

- predictions obtained from Pesaran and Timmermann's (2007) model-averaging procedure (BP-MA); if  $\tilde{\beta}_T$  is the BP-MA estimate of  $\beta_T$  (see above), then:

$$\tilde{y}_{T+1} = x_{T+1} \tilde{\beta}_T$$

The Monte Carlo replications are used to estimate the mean squared error of the

predictions:

$$MSE_j^y = \text{E} [(y_{T+1} - \tilde{y}_{T+1})^2]$$

where  $j = \text{TVC-MA, TVC-MS, TVC-II, TVC-}\pi, \text{OLS, BP, BP-MA}$  depending on which of the above methods has been used to forecast  $y_{T+1}$ .

To increase the accuracy of our Monte Carlo estimates of  $MSE_j^y$ , we use the fact that:

$$\text{E} [(y_{T+1} - \tilde{y}_{T+1})^2] = \text{E} [v_{T+1}^2] + \text{E} \left[ \left( \beta_{T+1} - \tilde{\beta}_{T+1} \right)^\top x_{T+1}^\top x_{T+1} \left( \beta_{T+1} - \tilde{\beta}_{T+1} \right) \right]$$

Since  $\text{E} [v_{T+1}^2]$  is known, we use the Monte Carlo simulations to estimate only the second summand on the right hand side of the above equation.

Table 2 reports the Monte Carlo estimates of  $MSE_j^y$ . The variation in  $MSE_j^y$  across models and design parameters broadly reflects the variation in  $MSE_j^\beta$  we have discussed above. To avoid repetitions, we point out the only significant difference, which concerns the highly persistent design ( $\rho = 0.99$ ): while the TVC and BP estimates give rise to an  $MSE_j^\beta$  that is around two orders of magnitude higher than  $MSE_{OLS}^\beta$ , the part of their  $MSE_j^y$  attributable to estimation error ( $MSE_j^y - 1$ ) compares much more favorably to its OLS counterpart, especially in the designs where  $x_t$  includes three lags of  $y_t$  and  $u_t$ . This might be considered evidence of the identification problem mentioned above.

## 5.2 Performance when the DGP is a regression with a discrete structural break

In this subsection we present the results of a set of Monte Carlo simulations aimed at understanding how our TVC model performs when regression coefficients experience a single discrete structural break. As in the previous subsection, we analyze both losses in forecasting performance and losses in estimation precision.

The Monte Carlo design is the same employed in the previous subsection, except for the fact that the data generating process is now subject to a discrete structural break at an unknown date:



- *Data generating process:*  $y_t$  is generated according to:

$$y_t = \rho y_{t-1} + u_{t-1} + v_t \text{ if } t < \tau$$

$$y_t = \rho y_{t-1} + (1 + b) u_{t-1} + v_t \text{ if } t \geq \tau$$

where  $y_0 = 0$ ,  $u_t \sim T(0, 1, 5)$  i.i.d.,  $v_t \sim N(0, 1)$  i.i.d. and  $u_t$  and  $v_t$  are serially and cross-sectionally independent;  $\tau$  is the stochastic breakpoint date, extracted from a discrete uniform distribution on the set of sample dates (from 1 to  $T$ );  $b \sim N(0, 1)$  is the stochastic break in regression coefficients.

The estimation precision and the forecasting performance are evaluated comparing the estimates of the coefficient vector at time  $T$  and the predictions of  $y_{T+1}$  with their true values.

Panel A of Table 3 reports the Monte Carlo estimates of  $MSE_j^\beta$  for the case in which  $x_t$  includes only the first lags of  $y_t$  and  $u_t$ . As before, we first discuss the cases in which  $\rho \neq 0.99$ . The OLS estimates, which have the smallest MSEs in the stable case (see previous subsections) are now those with the highest MSEs. Both the frequentist methods (BP and BP-MS) and the TVC methods (all four kinds) achieve a significant reduction of the MSE with respect to OLS. Although TVC-MA and TVC-MS perform slightly better than TVC-II and TVC- $\pi$ , there is not a clear ranking between the former two and the two frequentist methods: their MSEs are on average comparable, but TVC-MA and TVC-MS tend to perform better when the sample size is small ( $T = 100$ ), while BP and BP-MA tend to perform better when the sample size is large ( $T = 200, 500$ ). This might be explained by the fact that BP and BP-MA require the estimation of a considerable number of parameters when one or more break-dates are found and these parameters are inevitably estimated with low precision when the sample size is small. In the case in which  $\rho = 0.99$ , results are again substantially different: the MSEs of the TVC estimates (all four kinds) and of the BP estimates become much larger than the MSEs of the OLS estimates (and the BP estimates fare better than the TVC estimates), while the MSEs of the BP-MA estimates remain below those of the OLS estimates. The remarks about potential identification problems made in the previous subsections apply also to these results.

Panel B of Table 3 reports the Monte Carlo estimates of  $MSE_j^\beta$  for the case in which  $x_t$  includes three lags of  $y_t$  and  $u_t$ . The patterns are roughly the same found in Panel A (see the previous paragraph), with the relative performance of the TVC

methods and the frequentist methods depending on the sample size  $T$ . The only difference worth mentioning is that when  $\rho = 0.99$  the increase in the MSEs is milder and the TVC-MA estimates are more precise than the BP estimates.

As far as out-of-sample forecasting performance is concerned (Table 4, Panels A and B), the patterns in the  $MSE_j^y$  broadly reflect the patterns in the  $MSE_j^\beta$ . Again, there is an exception to this: when  $\rho = 0.99$ , high values of  $MSE_j^\beta$  do not translate into high values of  $MSE_j^y$ ; as a consequence, despite the aforementioned identification problem, the BP and the four TVC forecasts are much more accurate than the OLS forecasts (and in some cases also more accurate than the BP-MA forecasts).

### 5.3 Performance when the DGP is a regression with frequently changing coefficients

In this subsection we present the results of a set of Monte Carlo simulations aimed at understanding how our TVC model performs when regression coefficients experience frequent changes.

We analyze both losses in forecasting performance and losses in estimation precision, using the same Monte Carlo design employed in the previous two subsections. The only difference is that the data is now generated by a regression whose coefficients change at every time period:

- *Data generating process:*  $y_t$  is generated according to:

$$\begin{aligned} y_t &= \rho y_{t-1} + b_t u_{t-1} + v_t \\ b_t &= b_{t-1} + w_t \end{aligned}$$

where  $y_0 = 0$ ,  $b_0 = 1$ ,  $u_t \sim T(0, 1, 5)$  i.i.d.,  $v_t \sim N(0, 1)$  i.i.d.,  $w_t \sim N(0, W)$  i.i.d. and  $u_t$ ,  $v_t$  and  $w_t$  are serially and cross-sectionally independent. To ease comparisons with the previous subsection,  $W$  is chosen in such a way that  $b_T \sim N(1, 1)$ , irrespective of the sample size  $T$ :

$$W = \frac{1}{T}$$

Note that, although one coefficient of the regression is frequently changing ( $b_t$ ), the other coefficient ( $\rho$ ) is stable. As a consequence, the true DGP does not fit

exactly any of the possible DGPs contemplated by the TVC model. We prefer to adopt this specification over a specification in which the TVC model is correctly specified, because the results obtained with the latter specification are trivial (the TVC estimates are the best possible estimates). Furthermore, controlling  $\rho$  (keeping it fixed) allows to better understand its effects on model performance.

Panel A of Table 5 reports the Monte Carlo estimates of  $MSE_j^\beta$  for the case in which  $x_t$  includes only the first lags of  $y_t$  and  $u_t$ . We first summarize the results obtained when  $\rho \neq 0.99$ . The lowest MSEs are achieved by the TVC-MA estimates. The TVC-MS estimates are the second best (in some cases  $MSE_{TVC-MS}^\beta$  is almost identical to  $MSE_{TVC-MA}^\beta$ ). TVC-II and TVC- $\pi$  also have a performance comparable to that of TVC-MA (the increase in the MSEs is on average less than 5 per cent). The BP estimates are significantly less precise than the TVC estimates (their MSEs are roughly between 30 and 70 per cent higher than  $MSE_{TVC-MA}^\beta$ ). Finally, BP and BP-MA have a comparable performance when  $T = 100$ , but BP-MA is much less precise when the sample size increases ( $T = 200, 500$ ).

When  $\rho = 0.99$ , we again observe a sharp increase in the MSEs of the TVC estimates (all four kinds) and of the BP estimates: their MSEs become several times those of the OLS estimates. BP-MA achieves a significant reduction in MSE over OLS with larger sample sizes ( $T = 200, 500$ ). Thus, also with frequently changing coefficients, BP-MA seems to be the only method capable of dealing simultaneously with coefficient instability and a highly persistent lagged dependent variable.

Panel B of Table 5 reports the Monte Carlo estimates of  $MSE_j^\beta$  for the case in which  $x_t$  includes three lags of  $y_t$  and  $u_t$ . Similarly to what we found in the previous subsections, the only noticeable difference with respect to the one-lag case is that when  $\rho = 0.99$  the increase in the MSEs is milder.

As far as out-of-sample forecasting performance is concerned (Table 6, Panels A and B), the patterns in the  $MSE_j^y$  broadly reflect the patterns in the  $MSE_j^\beta$ . Again, the case  $\rho = 0.99$  constitutes an exception: despite their high  $MSE_j^\beta$ , the BP and the four TVC forecasts are more accurate than the OLS forecasts (and the TVC-MA and TVC-II forecasts are also more accurate than the BP-MA forecasts).

## 6 Empirical application: estimating common stocks' exposures to risk factors

In this section we briefly illustrate an empirical application of our TVC model. We use the model to estimate the exposures of S&P 500 constituents to market-wide risk factors. We track the weekly returns of the S&P 500 constituents for 10 years (from January 2000 to December 2009). An uninterrupted time series of returns is available for 432 of the 500 constituents (as of December 2009). The list of constituents and their returns are downloaded from Datastream. The risk factors we consider are the Fama and French's (1993, 1995 and 1996) risk factors (excess return on the market portfolio, return on the Small Minus Big portfolio, return on the High Minus Low portfolio), downloaded from Kenneth French's website.

The exposures to the risk factors are the coefficients  $\beta_t$  in the regression

$$y_t = x_t \beta_t + v_t$$

where  $y_t$  is the excess return on a stock at time  $t$ ,

$$x_t = \left[ 1 \quad r_{M,t} - r_{f,t} \quad SMB_t \quad HML_t \right]$$

$r_{M,t}$  is the return on the market portfolio at time  $t$ ,  $r_{f,t}$  is the risk-free rate of return and  $SMB_t$  and  $HML_t$  are the returns at time  $t$  on the SMB and HML portfolios respectively.

The procedures illustrated in the previous section are employed to understand whether the risk exposures  $\beta_t$  are time-varying and whether the TVC model provides good estimates of these risk exposures.

For a vast majority of the stocks included in our sample, we find evidence that  $\beta_t$  is indeed time-varying<sup>16</sup>.  $\theta = 0$  is the posterior mode of the mixing parameter only for 11 stocks out of 432. Furthermore,  $\Pi < 0.1$  and  $\pi < 0.1$  for 92% and 81% of the stocks respectively. On average,  $\Pi$  is 0.046 and  $\pi$  is 0.010. Also the frequentist method provides evidence that most stocks experience instability in their risk exposures: according to the BP sequential estimates, more than 78% of stocks experience at least one break in  $\beta_t$ .

---

<sup>16</sup>For example, the stock prices of the major US banks became more exposed to the return of the market portfolio during the last decade.

To evaluate the forecasting performance, we use the out-of-sample forecasts of  $y_t$  obtained after the first 400th week. The methods used to make predictions are those described in the previous section ( $j = \text{TVC-MA, TVC-MS, TVC-II, TVC-}\pi, \text{OLS, BP, BP-MA}$ ). For each stock  $i$  and for each prediction method  $j$ , the mean squared error is computed as:

$$\overline{MSE}_{i,j} = \frac{1}{T - T_0} \sum_{t=T_0+1}^T (y_{t,i} - \tilde{y}_{t,i,j})^2$$

where  $T_0$  is the number of periods elapsed before the first out-of-sample forecast is produced,  $\tilde{y}_{t,i,j}$  denotes the prediction of the excess return of the  $i$ -th stock at time  $t$ , conditional on  $x_t$ , produced by method  $j$ , and  $y_{t,i}$  is the corresponding realization.

To be able to compare the performance of the various methods across stocks, we use the performance of OLS forecasts as a benchmark. Thus, the gain from using model  $j$  with stock  $i$  is defined as:

$$GAIN_{i,j} = 1 - \frac{\overline{MSE}_{i,j}}{\overline{MSE}_{i,OLS}}$$

i.e.  $GAIN_{i,j}$  is the average reduction in MSE achieved by using model  $j$  instead of OLS. A positive value indicates an improvement in forecasting performance.

Table 7 reports some summary statistics of the sample distribution of  $GAIN_{i,j}$  (each stock  $i$  represents a sample point). All the TVC methods achieve a reduction in MSE and, among the TVC methods, TVC-MA achieves the maximum average reduction (approximately 3 per cent). BP performs very poorly (it actually causes a strong increase in MSE), while the average reduction achieved by BP-MA is similar to that of TVC-MA (again, approximately 3 per cent). The four TVC models have similar sample distributions of gains, characterized by a pronounced skew to the right (several small gains and few very large gains); furthermore, all four have a more dispersed distribution than the BP-MA model.

## 7 Conclusions

We have proposed a Bayesian regression model with time-varying coefficients (TVC). With respect to existing TVC models, we have introduced some technical innovations

aimed at making TVC models less computationally expensive and completely automatic (by completely automatic we mean that regressors and regressands are the only input required from the econometrician, so that he/she does not need to engage in technically demanding specifications of priors and model parametrizations).

We have conducted several Monte Carlo experiments to understand the pros and cons that might be encountered when using the TVC model in applied econometric analyses. We have found that the cons are generally limited, in the sense that the TVC model has satisfactory estimation precision and forecasting performance also when regression coefficients are indeed stable or when coefficient instability is present but the TVC model is mis-specified. In the presence of coefficient instability, there are potential rewards from using the TVC model: in some cases, its estimation precision and forecasting accuracy are significantly better than those of competing models.

To demonstrate a real-world application of our TVC model, we have used it to estimate the exposures of S&P 500 stocks to market-wide risk factors. We have found that a vast majority of stocks have time-varying risk exposures and that the TVC model helps to better forecast these exposures.

Before concluding, two remarks on the applicability of our TVC model are in order. First, we have confined attention to single equation regression models, but the results presented in the paper can be extended in a straightforward manner to multiple equation models (for example VARs), by imposing the usual normal / inverse Wishart priors on the initial parameters. Second, we have not discussed the use of the model for the analysis of cross-sectional data: however, it is possible to use TVC models like ours to analyze cross-sectional data in the presence of non-linearities that are not explicitly captured by the regressors (see West and Harrison - 1997); this is usually accomplished by replacing the time index  $t$  with the rank statistic of the regressor that is presumably responsible for the non-linearity.

## 8 Appendix

### 8.1 Proofs of propositions 2 and 3

In this section we derive the formulae presented in Propositions 2 and 3. To facilitate the exposition, we start from simpler information structures and then we tackle the more complex information structure assumed in Propositions 2 and 3 and summarized in Assumption 1.

#### 8.1.1 $V$ and $\theta$ known, $\beta_1$ unknown

We start from the simple case in which  $V$  and  $\theta$  are both known. The assumptions on the priors and the initial information are summarized as follows:

**Case 17 (Priors and initial information)** *The priors on the unknown parameters are:*

$$(\beta_1 | D_0^{**}) \sim N\left(\widehat{\beta}_{1,0}, VF_{\beta,1,0}\right)$$

and the initial information set is:

$$D_0^{**} = \left\{ \widehat{\beta}_{1,0}, F_{\beta,1,0}, V, \theta \right\}$$

Note that also  $W^* = \lambda(\theta) F_{\beta,1,0}$  and  $W = V\lambda(\theta) F_{\beta,1,0}$  are known, because  $\theta$  and  $V$  are known. The information sets  $D_t^{**}$  satisfy the recursion  $D_t^{**} = D_{t-1}^{**} \cup \{y_t, x_t\}$ , starting from the set  $D_0^{**}$ . Given the above assumptions, as new information becomes available, the posterior distribution of the parameters of the regression can be calculated using the following results:

**Proposition 18 (Forward updating)** *Let priors and initial information be as in Case 17. Then:*

$$\begin{aligned} (\beta_t | D_{t-1}^{**}) &\sim N\left(\widehat{\beta}_{t,t-1}, VF_{\beta,t,t-1}\right) \\ (y_t | D_{t-1}^{**}, x_t) &\sim N\left(\widehat{y}_{t,t-1}, VF_{y,t,t-1}\right) \\ (\beta_t | D_t^{**}) &\sim N\left(\widehat{\beta}_{t,t}, VF_{\beta,t,t}\right) \end{aligned}$$

where the means and variances of the above distributions are calculated recursively as

follows:

$$\begin{aligned}
\widehat{\beta}_{t,t-1} &= \widehat{\beta}_{t-1,t-1} & F_{\beta,t,t-1} &= F_{\beta,t-1,t-1} + W^* \\
\widehat{y}_{t,t-1} &= x_t \widehat{\beta}_{t,t-1} & F_{y,t,t-1} &= 1 + x_t F_{\beta,t,t-1} x_t^\top \\
e_t &= y_t - \widehat{y}_{t,t-1} & P_t &= F_{\beta,t,t-1} x_t^\top / F_{y,t,t-1} \\
\widehat{\beta}_{t,t} &= \widehat{\beta}_{t,t-1} + P_t e_t & F_{\beta,t,t} &= F_{\beta,t,t-1} - P_t P_t^\top F_{y,t,t-1}
\end{aligned} \tag{12}$$

starting from the initial values  $\widehat{\beta}_{1,0}$  and  $F_{\beta,1,0}$ .

**Proof.** Note that, given the above assumptions, the system:

$$\begin{cases} y_t = x_t \beta_t + v_t \\ \beta_t = \beta_{t-1} + w_t \end{cases}$$

is a Gaussian linear state-space system, where  $y_t = x_t \beta_t + v_t$  is the observation equation and  $\beta_t = \beta_{t-1} + w_t$  is the transition equation. Hence, the posterior distribution of the states can be updated using the Kalman filter. The recursive equations (12) are just the usual updating equations of the Kalman filter (e.g.: Hamilton - 1994). ■

The smoothing equations are provided by the following proposition:

**Proposition 19 (Backward updating)** *Let priors and initial information be as in Case 17. Then:*

$$(\beta_{T-\tau} | D_T^{**}) \sim N \left( \widehat{\beta}_{T-\tau,T}, V F_{\beta,T-\tau,T} \right)$$

where the means and the variances of the above distributions are calculated recursively (backwards) as follows:

$$\begin{aligned}
Q_{T-\tau} &= F_{\beta,T-\tau,T-\tau} (F_{\beta,T-\tau+1,T-\tau})^{-1} \\
\widehat{\beta}_{T-\tau,T} &= \widehat{\beta}_{T-\tau,T-\tau} + Q_{T-\tau} \left( \widehat{\beta}_{T-\tau+1,T} - \widehat{\beta}_{T-\tau+1,T-\tau} \right) \\
F_{\beta,T-\tau,T} &= F_{T-\tau,T-\tau} + Q_{T-\tau} (F_{T-\tau+1,T} - F_{T-\tau+1,T-\tau}) Q_{T-\tau}^\top
\end{aligned}$$

and the backward recursions start from the terminal values of the forward recursions (12).

**Proof.** These are the usual backward Kalman recursions (e.g.: Hamilton - 1994). ■

### 8.1.2 $\theta$ known, $\beta_1$ and $V$ unknown

In this subsection we relax the assumption that  $V$  (the variance of  $v_t$ ) is known and we impose a Gamma prior on the reciprocal of  $V$ . The assumptions on the priors and



the initial information are summarized as follows:

**Case 20 (Priors and initial information)** *The priors on the unknown parameters are:*

$$\begin{aligned}(\beta_1 | D_0^*) &\sim N\left(\widehat{\beta}_{1,0}, VF_{\beta,1,0}\right) \\(1/V | D_0^*) &\sim G\left(\widehat{V}_0, n_0\right)\end{aligned}$$

and the initial information set is:

$$D_0^* = \left\{ \widehat{\beta}_{1,0}, F_{\beta,1,0}, \widehat{V}_0, n_0, \theta \right\}$$

Note that also  $W^* = \lambda(\theta) F_{\beta,1,0}$  is known, because  $\theta$  is known. The information sets  $D_t^*$  satisfy the recursion  $D_t^* = D_{t-1}^* \cup \{y_t, x_t\}$ , starting from the set  $D_0^*$ . Given the above assumptions, the posterior distributions of the parameters of the regression can be calculated as follows:

**Proposition 21 (Forward updating)** *Let priors and initial information be as in Case 20. Then:*

$$\begin{aligned}(\beta_t | D_{t-1}^*) &\sim T\left(\widehat{\beta}_{t,t-1}, \widehat{V}_{t-1} F_{\beta,t,t-1}, n_{t-1}\right) \\(y_t | D_{t-1}^*, x_t) &\sim T\left(\widehat{y}_{t,t-1}, \widehat{V}_{t-1} F_{y,t,t-1}, n_{t-1}\right) \\(\beta_t | D_t^*) &\sim T\left(\widehat{\beta}_{t,t}, \widehat{V}_t F_{\beta,t,t}, n_t\right) \\(1/V | D_t^*) &\sim G\left(\widehat{V}_t, n_t\right)\end{aligned}$$

where the parameters of the above distributions are calculated recursively as in (12) and as follows:

$$\begin{aligned}n_t &= n_{t-1} + 1 \\ \widehat{V}_t &= \frac{1}{n_t} \left( n_{t-1} \widehat{V}_{t-1} + \frac{e_t^2}{F_{y,t,t-1}} \right)\end{aligned}$$

starting from the initial values  $\widehat{\beta}_{1,0}$ ,  $F_{\beta,1,0}$ ,  $\widehat{V}_0$  and  $n_0$ .

**Proof.** The proof is by induction. At time  $t = 1$ ,  $p(\beta_1 | D_0^*, V)$  and  $p(1/V | D_0^*)$  are the conjugate normal / inverse gamma priors of a standard Bayesian regression model

with constant coefficients (e.g. Hamilton - 1994). Therefore, the usual results on the updating of these conjugate priors hold:

$$(\beta_1 | D_1^*, V) \sim N(\widehat{\beta}_{1,1}, VF_{\beta,1,1}) \quad (13)$$

$$(1/V | D_1^*) \sim G(\widehat{V}_1, n_1) \quad (14)$$

Since  $\beta_2 = \beta_1 + w_2$  and

$$(w_2 | D_1^*, V) \sim N(0, VW^*)$$

then, by the additivity of normal distributions:

$$(\beta_2 | D_1^*, V) \sim N(\widehat{\beta}_{1,1}, VF_{\beta,1,1} + VW^*) = N(\widehat{\beta}_{2,1}, VF_{\beta,2,1})$$

Therefore, at time  $t = 2$ ,  $p(\beta_2 | D_1^*, V)$  and  $p(1/V | D_1^*)$  are again the conjugate normal / inverse gamma priors of a standard Bayesian regression model with constant coefficients. Proceeding in the same way as for  $t = 1$ , one obtains the desired result for  $t = 2$  and, inductively, for all the other periods. ■

Posterior distributions of the coefficients that take into account all information received up to time  $T$  are calculated as follows:

**Proposition 22 (Backward updating)** *Let priors and initial information be as in Case 20. Then:*

$$(\beta_{T-\tau} | D_T^*) \sim T(\widehat{\beta}_{T-\tau,T}, \widehat{V}_T F_{\beta,T-\tau,T}, n_T)$$

where  $\widehat{V}_T$  and  $n_T$  are calculated as in Proposition 18 and the other parameters of the above distributions are calculated recursively (backwards) as in Proposition 19.

**Proof.** From Proposition 19, we know that:

$$(\beta_{T-\tau} | D_T^*, V) = (\beta_{T-\tau} | D_T^{**}) \sim N(\widehat{\beta}_{T-\tau,T}, VF_{\beta,T-\tau,T})$$

Furthermore,  $(1/V | D_T^*) \sim G(\widehat{V}_T, n_T)$ . By the conjugacy of  $(\beta_{T-\tau} | D_T^*, V)$  and  $(1/V | D_T^*)$ , it follows that:

$$(\beta_{T-\tau} | D_T^*) \sim T(\widehat{\beta}_{T-\tau,T}, \widehat{V}_T F_{\beta,T-\tau,T}, n_T)$$

■

### 8.1.3 $\theta$ , $\beta_1$ and $V$ unknown

In this subsection we relax the assumption that  $\theta$  is known, using the same priors and initial information of the propositions in the main text of the article (Propositions 2 and 3):

**Case 23** *The priors on the unknown parameters are:*

$$\begin{aligned} (\beta_1 | D_0, V, \theta) &\sim N\left(\widehat{\beta}_{1,0}, VF_{\beta,1,0}\right) \\ (1/V | D_0, \theta) &\sim G\left(\widehat{V}_0, n_0\right) \\ p(\theta_i | D_0) &= p_{0,i}, i = 1, \dots, q \end{aligned}$$

and the initial information set is:

$$D_0 = \left\{ \widehat{\beta}_{1,0}, F_{\beta,1,0}, \widehat{V}_0, n_0, p_{0,1}, \dots, p_{0,q} \right\}$$

The information sets  $D_t$  satisfy the recursion  $D_t = D_{t-1} \cup \{y_t, x_t\}$ , starting from the set  $D_0$ . Note that the assumptions introduced in Cases 17 and 20 in the previous subsections had the only purpose of introducing the more complex Case 23. Given the above assumptions, the posterior distributions of the parameters of the regression can be calculated as follows:

**Proposition 24** *Let priors and initial information be as in Case 23. Let  $p_{t,i} = p(\theta = \theta_i | D_t)$ . Then:*

$$\begin{aligned} p(\beta_t | D_{t-1}) &= \sum_{i=1}^q p(\beta_t | \theta = \theta_i, D_{t-1}) p_{t-1,i} \\ p(y_t | D_{t-1}, x_t) &= \sum_{i=1}^q p(y_t | \theta = \theta_i, D_{t-1}, x_t) p_{t-1,i} \\ p(1/V | D_{t-1}) &= \sum_{i=1}^q p(1/V | \theta = \theta_i, D_{t-1}) p_{t-1,i} \\ p(\beta_t | D_t) &= \sum_{i=1}^q p(\beta_t | \theta = \theta_i, D_t) p_{t,i} \end{aligned}$$

The mixing probabilities are obtained recursively as:

$$p_{t,i} = \frac{p_{t-1,i} p(y_t | \theta = \theta_i, D_{t-1}, x_t)}{\sum_{j=1}^q p_{t-1,j} p(y_t | \theta = \theta_j, D_{t-1}, x_t)}$$

starting from the prior probabilities  $p_{0,1}, \dots, p_{0,q}$ . The conditional densities

$$\begin{aligned} p(\beta_t | D_{t-1}, \theta = \theta_i) &= p(\beta_t | D_{t-1}^*) \\ p(y_t | D_{t-1}, x_t, \theta = \theta_i) &= p(y_t | D_{t-1}^*, x_t) \\ p(1/V | D_{t-1}, \theta = \theta_i) &= p(1/V | D_{t-1}^*) \\ p(\beta_t | D_t, \theta = \theta_i) &= p(\beta_t | D_t^*) \end{aligned}$$

are calculated for each  $\theta_i$  as in Propositions 18 and 21.

**Proof.** Conditioning on  $\theta = \theta_i$ , the distributions of the parameters  $\beta_t$  and  $V$  and of the observations  $y_t$  are obtained from Propositions 18 and 21 (it suffices to note that  $D_t \cup \theta = D_t^*$ ). Not conditioning on  $\theta = \theta_i$ , the distributions of the parameters  $\beta_t$  and  $V$  and of the observations  $y_t$  are obtained marginalizing their joint distribution with  $\theta$ . For example:

$$\begin{aligned} p(\beta_t | D_{t-1}) &= \sum_{i=1}^q p(\beta_t, \theta_i | D_{t-1}) \\ &= \sum_{i=1}^q p(\beta_t | \theta = \theta_i, D_{t-1}) p(\theta = \theta_i | D_{t-1}) \\ &= \sum_{i=1}^q p(\beta_t | \theta = \theta_i, D_{t-1}) p_{t-1,i} \end{aligned}$$

The mixing probabilities are obtained using Bayes' rule:

$$\begin{aligned}
p_{t,i} &= p(\theta = \theta_i | D_t) \\
&= p(\theta = \theta_i | y_t, D_{t-1}, x_t) \\
&= \frac{p(y_t | \theta = \theta_i, D_{t-1}, x_t) p(\theta = \theta_i | D_{t-1}, x_t)}{p(y_t | D_{t-1}, x_t)} \\
&= \frac{p(y_t | \theta = \theta_i, D_{t-1}, x_t) p(\theta = \theta_i | D_{t-1})}{\sum_{j=1}^q p(y_t, \theta_j | D_{t-1}, x_t)} \\
&= \frac{p(y_t | \theta = \theta_i, D_{t-1}, x_t) p(\theta = \theta_i | D_{t-1})}{\sum_{j=1}^q p(y_t | \theta = \theta_j, D_{t-1}, x_t) p(\theta = \theta_j | D_{t-1}, x_t)} \\
&= \frac{p(y_t | \theta = \theta_i, D_{t-1}, x_t) p_{t-1,i}}{\sum_{j=1}^q p(y_t | \theta = \theta_j, D_{t-1}, x_t) p_{t-1,j}}
\end{aligned}$$

■

Proposition 2 in the main text is obtained by combining propositions 18, 21 and 24 above. Proposition 3 results from propositions 19, 22 and 24 above.

## 8.2 Proof of proposition 14 (scale invariance)

When  $x_t R$  is the vector of regressors, the prior covariance is:

$$F_{\beta,1,0} = T (R^\top X^\top X R)^{-1} = T R^{-1} (X^\top X)^{-1} (R^{-1})^\top$$

The constant  $\omega$  is unaffected by the rotation:

$$\begin{aligned}
\omega &= \frac{1}{T} \sum_{t=1}^T x_t R F_{\beta,1,0} R^\top x_t^\top \\
&= \frac{1}{T} \sum_{t=1}^T x_t R \left( T R^{-1} (X^\top X)^{-1} (R^{-1})^\top \right) R^\top x_t^\top \\
&= \frac{1}{T} \sum_{t=1}^T x_t \left( T (X^\top X)^{-1} \right) x_t^\top
\end{aligned}$$

Since  $R_\theta$  does not depend on the data and

$$\lambda(\theta) = \frac{\theta}{\omega - \theta\omega}$$

the fact that  $\omega$  does not change implies that also  $R_\lambda$  (the support of  $\lambda$ ) remains unchanged. The prior probabilities assigned to the elements of  $R_\lambda$  also do not depend on the data. So, the prior distribution of  $\lambda$  is unaffected by the rotation.

As far as the recursive equations in Proposition 2 are concerned, note that the initial conditions

$$\begin{aligned}\widehat{\beta}_{1,0,i} &= \widehat{\beta}_{1,0} = 0 \\ \widehat{V}_{0,i} &= \widehat{V}_0 = y_0^2 \\ n_{0,i} &= n_0 = 1\end{aligned}$$

are not affected by the rotation, while the initial condition

$$F_{\beta,1,0,i} = F_{\beta,1,0} = TR^{-1} (X^\top X)^{-1} (R^{-1})^\top$$

changes (it is pre-multiplied by  $R^{-1}$  and post-multiplied by  $(R^{-1})^\top$ ).

For  $t > 0$ , it can be easily checked that  $\widehat{y}_{t,t-1,i}$ ,  $F_{y,t,t-1,i}$ ,  $e_{t,i}$ ,  $n_{t,i}$ ,  $\widehat{V}_{t,i}$  remain unchanged, while  $F_{\beta,t,t-1,i}$  and  $F_{\beta,t,t,i}$  are pre-multiplied by  $R^{-1}$  and post-multiplied by  $(R^{-1})^\top$  and  $\widehat{\beta}_{t,t-1,i}$ ,  $\widehat{\beta}_{t,t,i}$  and  $P_{t,i}$  are pre-multiplied by  $R^{-1}$ . Therefore:

$$(\beta_t | D_{t-1}^*, \lambda = \lambda(\theta_i)) = (R^{-1}\beta_t | D_{t-1}, \lambda = \lambda(\theta_i)) \quad \forall t \leq T, i = 1, \dots, q$$

and

$$(\beta_t | D_t^*, \lambda = \lambda(\theta_i)) = (R^{-1}\beta_t | D_t, \lambda = \lambda(\theta_i)) \quad \forall t \leq T, i = 1, \dots, q$$

The model probabilities  $p_{t,1}, \dots, p_{t,q}$  depend only on  $\widehat{y}_{t,t-1,i}$ ,  $F_{y,t,t-1,i}$ ,  $n_{t-1,i}$  and  $\widehat{V}_{t-1,i}$ , which remain unchanged, so they remain unchanged as well. As a consequence, also unconditionally:

$$(\beta_t | D_s^*) = (R^{-1}\beta_t | D_s) \quad \forall t \leq T$$

for  $s = t - 1$  or  $s = t$ . Using similar arguments on the backward recursions of proposition 3, it is possible to prove that the above equality holds for any  $s \leq T$ .

## References

- [1] Andrews, D. W. K., I. Lee and W. Ploberger (1996) "Optimal changepoint tests for normal linear regression model", *Journal of Econometrics*, 70, 9-38.
- [2] Bai, J. and P. Perron (1998) "Estimating and testing linear models with multiple structural breaks", *Econometrica*, 66, 47-78.
- [3] Bai, J. and P. Perron (2003) "Computation and analysis of multiple structural change models", *Journal of Applied Econometrics*, 18, 1-22.
- [4] Bai, J. and P. Perron (2006) "Multiple structural change models: a simulation analysis", *Econometric theory and practice: frontier of analysis and applied research (Essays in honor of Peter Phillips)* ed. by Corbae, D., S. Durlauf and B. E. Hansen, Cambridge University Press.
- [5] Brown, R., L., J. Durbin and J. M. Evans (1975) "Techniques for testing the constancy of regression coefficients over time", *Journal of the Royal Statistical Society B*, 37, 149-192.
- [6] Canova, F. and M. Ciccarelli (2009) "Estimating multi-country VAR models", *International Economic Review*, 929-959.
- [7] Canova, F. and L. Gambetti (2009) "Structural changes in the US economy: is there a role for monetary policy?", *Journal of Economic Dynamics and Control*, 477-490.
- [8] Carter, C. K. and R. Kohn (1994) "On Gibbs sampling for state space models", *Biometrika*, 81, 541-553.
- [9] Chib, S. and E. Greenberg (1995) "Hierarchical analysis of SUR models with extensions to correlated serial errors and time-varying parameter models", *Journal of Econometrics*, 68, 339-360.
- [10] Chow, G. C. (1960) "Tests of equality between sets of coefficients in two linear regressions", *Econometrica*, 28, 591-605.
- [11] Cogley, T., G. E. Primiceri and T. J. Sargent (2010) "Inflation-gap persistence in the US", *American Economic Journal: Macroeconomics*, 43-69.

- [12] Cogley, T. and T.J. Sargent (2001) "Evolving Post World War II U.S. Inflation Dynamics", *NBER Macroeconomics Annual*, 16, 331-373.
- [13] Datta, G. S. (1996) "On priors providing frequentist validity of Bayesian inference for multiple parametric functions", *Biometrika*, 83, 287-298.
- [14] Datta, G. S. and J. K. Ghosh (1995) "On priors providing frequentist validity for Bayesian inference", *Biometrika*, 82, 37-45.
- [15] Doan, T., R. B. Litterman and C. A. Sims (1984) "Forecasting and conditional projection using realistic prior distributions", *Econometric Reviews*, 3, 1-100.
- [16] Fama, E. and K. French (1993) "Common risk factors in the returns on bonds and stocks", *Journal of Financial Economics*, 33, 3-56.
- [17] Fama, E. and K. French (1995) "Size and book-to-market factors in earnings and returns", *Journal of Finance*, 50, 131-155.
- [18] Fama, E. and K. French (1996) "Multifactor explanations of asset pricing anomalies", *Journal of Finance*, 51, 55-84.
- [19] Fernandez, C., E. Ley and M. F. Steel (2001) "Benchmark priors for Bayesian model averaging", *Journal of Econometrics*, 100, 381-427.
- [20] George, E. I. and R. E. McCulloch (1997) "Approaches for Bayesian variable selection", *Statistica Sinica*, 7, 339-374.
- [21] Ghosh, J. K. and R. Mukerjee (1993) "Frequentist validity of highest posterior density regions in the multiparameter case", *Annals of the Institute of Statistical Mathematics*, 45, 293-302.
- [22] Guerre, E. and P. Lavergne (2005) "Data-driven rate-optimal specification testing in regression models", *The Annals of Statistics*, 33, 840-870.
- [23] Hamilton J. D. (1994) *Time Series Analysis*, Princeton University Press, Princeton, USA.
- [24] Hansen, B. E. (2000) "Testing for structural change in conditional models", *Journal of Econometrics*, 97, 93-115.



- [25] Hatanaka, M. and K. Yamada (1999) "A unit root test in the presence of structural changes in I(1) and I(0) models", *Cointegration, causality and forecasting: a Festschrift in honour of Clive W. J. Granger* ed. by Engle R. F. and H. White, Oxford University Press.
- [26] Horowitz, J. L. and V. G. Spokoiny (2001) "An adaptive, rate-optimal test of a parametric model against a nonparametric alternative", *Econometrica*, 69, 599-631.
- [27] Jeffreys, H. (1961) *The theory of probability* (3e), Oxford University Press.
- [28] Kapetanios, G. (2008) "Bootstrap-based tests for deterministic time-varying coefficients in regression models", *Computational Statistics & Data Analysis*, 15, 534-545.
- [29] Kass, R. E. and L. Wasserman (1995) "A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion", *Journal of the American Statistical Association*, 90, 928-934.
- [30] Koop, G., R. Leon-Gonzalez, and R. Strachan (2009) "On the evolution of the monetary policy transmission mechanism", *Journal of Economic Dynamics and Control*, 997-1017.
- [31] Koop, G. and D. Korobilis (2009) "Bayesian multivariate time series methods for empirical macroeconomics", *Foundations and trends in econometrics*, 267-358.
- [32] Lepski, O. V., E. Mammen and V.G. Spokoiny (1997) "Optimal spatial adaptation to inhomogeneous smoothness: an approach based on kernel estimates with variable bandwidth selectors", *Annals of Statistics*, 25, 929-947.
- [33] Liang F., R. Paulo, G. Molina, M. A. Clyde and J. O. Berger (2008) "Mixtures of g priors for Bayesian variable selection", *Journal of the American Statistical Association*, 103, 410-423.
- [34] Mukerjee, R. and D. K. Dey (1993) "Frequentist validity of posterior quantiles in the presence of a nuisance parameter: higher order asymptotics", *Biometrika*, 80, 499-505.

- [35] Nyblom, J. (1989): "Testing for the constancy of parameters over time", *Journal of the American Statistical Association*, 84, 348-368.
- [36] Perron, P. and X. Zhu (2005) "Structural breaks with deterministic and stochastic trends", *Journal of Econometrics*, 129, 65-119.
- [37] Robert, C. P. (2007) *The Bayesian choice: from decision-theoretic foundations to computational implementation*, Springer Verlag.
- [38] Sargent, T., Williams, N. and T. Zha (2006) "Shocks and government beliefs: the rise and fall of American inflation", *American Economic Review*, 1193-1224.
- [39] Shively, T. S., R. Kohn and S. Wood (1999) "Variable selection and function estimation in additive nonparametric regression using a data-based prior (with discussion)", *Journal of the American Statistical Association*, 94, 777-806.
- [40] Smith, M. and R. Kohn (1996) "Nonparametric regression using Bayesian variable selection", *Journal of Econometrics*, 75, 317-343.
- [41] Spokoiny, V.G. (2001) "Data-driven testing the fit of linear models", *Mathematical Methods of Statistics*, 10, 465-497.
- [42] Stock, J. H. and M. W. Watson (1996) "Evidence on structural instability in macroeconomic time series relations", *Journal of Business & Economic Statistics*, 14, 11-30.
- [43] West, M. and J. Harrison (1997), *Bayesian forecasting and dynamic models*, Second Edition, Springer Verlag, New York.
- [44] Zellner, A. (1986) "On assessing prior distributions and Bayesian regression analysis with g-prior distributions", *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, eds. P. K. Goel and A. Zellner, 233-243, North-Holland/Elsevier.

## Tables

This section gathers all the tables described in the paper. All estimates of population quantities obtained from Monte Carlo simulations are complemented by an estimate of the Monte Carlo standard error (in parentheses).

**Table 1 – Estimation precision when coefficients are stable - Monte Carlo evidence - MSE of coefficient estimates**

**Panel A - One lag in the estimated equation**

	TVC-MA	TVC-MS	TVC- $II$	TVC- $\pi$	OLS	BP	BP-MA
	$\rho=0$						
T=100	0.0257 (0.0003)	0.0286 (0.0003)	0.0239 (0.0003)	0.0215 (0.0002)	0.0205 (0.0002)	0.0540 (0.0031)	0.0221 (0.0002)
T=200	0.0138 (0.0001)	0.0142 (0.0002)	0.0121 (0.0001)	0.0106 (0.0001)	0.0102 (0.0001)	0.0479 (0.0150)	0.0108 (0.0001)
T=500	0.0066 (0.0001)	0.0056 (0.0001)	0.0049 (0.0001)	0.0043 (0.0000)	0.0040 (0.0000)	0.0070 (0.0006)	0.0043 (0.0002)
	$\rho=0.50$						
T=100	0.0271 (0.0003)	0.0303 (0.0004)	0.0249 (0.0003)	0.0219 (0.0003)	0.0207 (0.0002)	0.0730 (0.0050)	0.0227 (0.0003)
T=200	0.0136 (0.0001)	0.0141 (0.0002)	0.0119 (0.0001)	0.0104 (0.0001)	0.0099 (0.0001)	0.0329 (0.0029)	0.0106 (0.0002)
T=500	0.0064 (0.0001)	0.0055 (0.0001)	0.0048 (0.0001)	0.0041 (0.0000)	0.0038 (0.0000)	0.0132 (0.0034)	0.0039 (0.0000)
	$\rho=0.80$						
T=100	0.0308 (0.0004)	0.0352 (0.0006)	0.0278 (0.0004)	0.0234 (0.0003)	0.0219 (0.0003)	0.1537 (0.0241)	0.0272 (0.0020)
T=200	0.0141 (0.0002)	0.0145 (0.0002)	0.0121 (0.0002)	0.0103 (0.0001)	0.0097 (0.0001)	0.0853 (0.0163)	0.0105 (0.0002)
T=500	0.0062 (0.0001)	0.0053 (0.0001)	0.0046 (0.0001)	0.0039 (0.0000)	0.0037 (0.0000)	0.0179 (0.0035)	0.0038 (0.0000)
	$\rho=0.99$						
T=100	4.7933 (0.1551)	5.3498 (0.1669)	4.7233 (0.1550)	4.7344 (0.1553)	0.0666 (0.0013)	1.9222 (0.3944)	0.1145 (0.0115)
T=200	2.5390 (0.0970)	2.7297 (0.1010)	2.5072 (0.0970)	2.5057 (0.0970)	0.0244 (0.0005)	0.7189 (0.1093)	0.0367 (0.0032)
T=500	0.4448 (0.0249)	0.4792 (0.0259)	0.4301 (0.0249)	0.4287 (0.0249)	0.0062 (0.0001)	0.1338 (0.0195)	0.0072 (0.0002)

**Panel B – Three lags in the estimated equation**

	TVC-MA	TVC-MS	TVC- $II$	TVC- $\pi$	OLS	BP	BP-MA
	$\rho=0$						
T=100	0.0885 (0.0007)	0.0964 (0.0009)	0.0860 (0.0007)	0.0804 (0.0006)	0.0785 (0.0006)	0.0827 (0.0013)	0.0795 (0.0006)
T=200	0.0433 (0.0003)	0.0453 (0.0004)	0.0414 (0.0003)	0.0385 (0.0003)	0.0376 (0.0003)	0.0379 (0.0003)	0.0378 (0.0003)
T=500	0.0186 (0.0001)	0.0175 (0.0001)	0.0162 (0.0001)	0.0150 (0.0001)	0.0146 (0.0001)	0.0146 (0.0001)	0.0146 (0.0001)
	$\rho=0.50$						
T=100	0.0953 (0.0008)	0.1033 (0.0009)	0.0920 (0.0007)	0.0863 (0.0007)	0.0842 (0.0006)	0.0905 (0.0021)	0.0850 (0.0006)
T=200	0.0457 (0.0004)	0.0478 (0.0004)	0.0434 (0.0003)	0.0406 (0.0003)	0.0398 (0.0003)	0.0399 (0.0003)	0.0398 (0.0003)
T=500	0.0194 (0.0002)	0.0184 (0.0002)	0.0170 (0.0001)	0.0159 (0.0001)	0.0155 (0.0001)	0.0157 (0.0002)	0.0156 (0.0001)
	$\rho=0.80$						
T=100	0.1049 (0.0009)	0.1134 (0.0010)	0.1017 (0.0008)	0.0959 (0.0008)	0.0940 (0.0007)	0.1024 (0.0041)	0.0947 (0.0007)
T=200	0.0511 (0.0004)	0.0532 (0.0005)	0.0484 (0.0004)	0.0454 (0.0004)	0.0442 (0.0003)	0.0447 (0.0004)	0.0443 (0.0003)
T=500	0.0215 (0.0002)	0.0203 (0.0002)	0.0188 (0.0002)	0.0174 (0.0001)	0.0169 (0.0001)	0.0169 (0.0001)	0.0169 (0.0001)
	$\rho=0.99$						
T=100	0.7213 (0.0409)	0.9599 (0.0498)	0.6244 (0.0402)	0.6713 (0.0408)	0.1505 (0.0016)	0.1830 (0.0072)	0.1531 (0.0019)
T=200	0.2619 (0.0203)	0.3052 (0.0226)	0.2317 (0.0199)	0.2425 (0.0202)	0.0632 (0.0006)	0.0867 (0.0142)	0.0636 (0.0007)
T=500	0.0327 (0.0017)	0.0310 (0.0018)	0.0279 (0.0017)	0.0264 (0.0017)	0.0213 (0.0002)	0.0213 (0.0002)	0.0213 (0.0002)

**Table 2 – Prediction accuracy when coefficients are stable - Monte Carlo evidence - MSE of one-step-ahead predictions**

**Panel A - One lag in the estimated equation**

	TVC-MA	TVC-MS	TVC- $II$	TVC- $\pi$	OLS	BP	BP-MA
$\rho=0$							
T=100	1.0384 (0.0010)	1.0438 (0.0013)	1.0358 (0.0009)	1.0319 (0.0008)	1.0301 (0.0006)	1.0942 (0.0098)	1.0332 (0.0008)
T=200	1.0208 (0.0005)	1.0214 (0.0005)	1.0181 (0.0004)	1.0161 (0.0004)	1.0155 (0.0004)	1.2914 (0.2190)	1.0165 (0.0004)
T=500	1.0101 (0.0002)	1.0084 (0.0002)	1.0073 (0.0002)	1.0064 (0.0002)	1.0060 (0.0001)	1.0121 (0.0026)	1.0073 (0.0012)
$\rho=0.50$							
T=100	1.0393 (0.0009)	1.0443 (0.0011)	1.0366 (0.0009)	1.0325 (0.0008)	1.0311 (0.0007)	1.0978 (0.0126)	1.0351 (0.0013)
T=200	1.0217 (0.0006)	1.0229 (0.0007)	1.0193 (0.0006)	1.0170 (0.0005)	1.0159 (0.0004)	1.0443 (0.0047)	1.0168 (0.0004)
T=500	1.0101 (0.0003)	1.0086 (0.0002)	1.0075 (0.0002)	1.0065 (0.0002)	1.0060 (0.0001)	1.0144 (0.0027)	1.0061 (0.0001)
$\rho=0.80$							
T=100	1.0453 (0.0011)	1.0526 (0.0014)	1.0423 (0.0011)	1.0371 (0.0010)	1.0348 (0.0009)	1.1301 (0.0212)	1.0420 (0.0021)
T=200	1.0218 (0.0006)	1.0230 (0.0008)	1.0191 (0.0006)	1.0162 (0.0003)	1.0155 (0.0003)	1.0757 (0.0212)	1.0166 (0.0005)
T=500	1.0103 (0.0003)	1.0089 (0.0003)	1.0078 (0.0002)	1.0065 (0.0002)	1.0062 (0.0001)	1.0229 (0.0084)	1.0063 (0.0001)
$\rho=0.99$							
T=100	1.2256 (0.0089)	1.2710 (0.0102)	1.2238 (0.0089)	1.2206 (0.0089)	1.0489 (0.0010)	1.2199 (0.0630)	1.0484 (0.0015)
T=200	1.1349 (0.0046)	1.1548 (0.0051)	1.1317 (0.0046)	1.1306 (0.0046)	1.0224 (0.0005)	1.0496 (0.0030)	1.0225 (0.0008)
T=500	1.0419 (0.0018)	1.0462 (0.0019)	1.0392 (0.0018)	1.0377 (0.0018)	1.0078 (0.0002)	1.0198 (0.0032)	1.0078 (0.0002)

**Panel B – Three lags in the estimated equation**

	TVC-MA	TVC-MS	TVC- $II$	TVC- $\pi$	OLS	BP	BP-MA
$\rho=0$							
T=100	1.0841 (0.0016)	1.0914 (0.0018)	1.0813 (0.0015)	1.0762 (0.0014)	1.0746 (0.0014)	1.0808 (0.0032)	1.0759 (0.0014)
T=200	1.0433 (0.0010)	1.0454 (0.0011)	1.0412 (0.0009)	1.0386 (0.0008)	1.0375 (0.0008)	1.0379 (0.0008)	1.0379 (0.0008)
T=500	1.0183 (0.0003)	1.0171 (0.0003)	1.0158 (0.0003)	1.0146 (0.0003)	1.0142 (0.0003)	1.0143 (0.0003)	1.0143 (0.0003)
$\rho=0.50$							
T=100	1.0845 (0.0016)	1.0923 (0.0019)	1.0817 (0.0016)	1.0769 (0.0014)	1.0752 (0.0014)	1.0928 (0.0119)	1.0772 (0.0015)
T=200	1.0435 (0.0008)	1.0454 (0.0009)	1.0412 (0.0008)	1.0380 (0.0007)	1.0372 (0.0007)	1.0375 (0.0007)	1.0375 (0.0007)
T=500	1.0179 (0.0003)	1.0167 (0.0003)	1.0154 (0.0003)	1.0144 (0.0003)	1.0141 (0.0003)	1.0141 (0.0003)	1.0141 (0.0003)
$\rho=0.80$							
T=100	1.0889 (0.0017)	1.0967 (0.0019)	1.0864 (0.0016)	1.0816 (0.0015)	1.0802 (0.0014)	1.0855 (0.0024)	1.0815 (0.0015)
T=200	1.0418 (0.0008)	1.0435 (0.0009)	1.0397 (0.0008)	1.0373 (0.0007)	1.0363 (0.0007)	1.0369 (0.0007)	1.0367 (0.0007)
T=500	1.0177 (0.0003)	1.0168 (0.0003)	1.0155 (0.0003)	1.0143 (0.0003)	1.0139 (0.0002)	1.0140 (0.0002)	1.0140 (0.0002)
$\rho=0.99$							
T=100	1.1381 (0.0065)	1.1705 (0.0080)	1.1298 (0.0062)	1.1353 (0.0064)	1.0943 (0.0017)	1.1128 (0.0145)	1.0886 (0.0016)
T=200	1.0647 (0.0031)	1.0727 (0.0038)	1.0609 (0.0030)	1.0617 (0.0030)	1.0439 (0.0008)	1.0504 (0.0075)	1.0425 (0.0008)
T=500	1.0196 (0.0004)	1.0192 (0.0005)	1.0178 (0.0004)	1.0171 (0.0004)	1.0163 (0.0003)	1.0160 (0.0003)	1.0160 (0.0003)

**Table 3 – Estimation precision when coefficients experience one discrete break - Monte Carlo evidence - MSE of coefficient estimates**

**Panel A - One lag in the estimated equation**

	TVC-MA	TVC-MS	TVC- $\Pi$	TVC- $\pi$	OLS	BP	BP-MA
	$\rho=0$						
T=100	0.1772 (0.0051)	0.1767 (0.0050)	0.1807 (0.0051)	0.1943 (0.0054)	0.3769 (0.0078)	0.1737 (0.0060)	0.2252 (0.0055)
T=200	0.1148 (0.0033)	0.1145 (0.0033)	0.1183 (0.0034)	0.1250 (0.0036)	0.3434 (0.0068)	0.1088 (0.0068)	0.1863 (0.0043)
T=500	0.0689 (0.0026)	0.0692 (0.0026)	0.0711 (0.0026)	0.0741 (0.0027)	0.3453 (0.0072)	0.0452 (0.0049)	0.1738 (0.0043)
	$\rho=0.50$						
T=100	0.1790 (0.0046)	0.1792 (0.0046)	0.1831 (0.0047)	0.1951 (0.0049)	0.3655 (0.0073)	0.2140 (0.0134)	0.2248 (0.0052)
T=200	0.1237 (0.0040)	0.1228 (0.0040)	0.1265 (0.0041)	0.1346 (0.0043)	0.3545 (0.0076)	0.1053 (0.0047)	0.1930 (0.0049)
T=500	0.0692 (0.0026)	0.0694 (0.0026)	0.0713 (0.0026)	0.0744 (0.0027)	0.3452 (0.0072)	0.0533 (0.0079)	0.1733 (0.0043)
	$\rho=0.80$						
T=100	0.1887 (0.0044)	0.1906 (0.0044)	0.1924 (0.0045)	0.2038 (0.0047)	0.3705 (0.0071)	0.3016 (0.0206)	0.2226 (0.0050)
T=200	0.1261 (0.0037)	0.1255 (0.0037)	0.1290 (0.0038)	0.1360 (0.0040)	0.3463 (0.0070)	0.1309 (0.0065)	0.1894 (0.0046)
T=500	0.0707 (0.0025)	0.0708 (0.0026)	0.0729 (0.0026)	0.0760 (0.0027)	0.3449 (0.0072)	0.0803 (0.0192)	0.1733 (0.0043)
	$\rho=0.99$						
T=100	5.1868 (0.1677)	5.7741 (0.1828)	5.1508 (0.1677)	5.1618 (0.1677)	0.4283 (0.0075)	2.5268 (0.2692)	0.4261 (0.0130)
T=200	3.3069 (0.1269)	3.5054 (0.1341)	3.2908 (0.1268)	3.2956 (0.1269)	0.3626 (0.0070)	1.4681 (0.1731)	0.2585 (0.0063)
T=500	0.7054 (0.0257)	0.7362 (0.0268)	0.7032 (0.0257)	0.7053 (0.0257)	0.3542 (0.0073)	0.6431 (0.1556)	0.1946 (0.0074)

**Panel B – Three lags in the estimated equation**

	TVC-MA	TVC-MS	TVC- $\Pi$	TVC- $\pi$	OLS	BP	BP-MA
	$\rho=0$						
T=100	0.3299 (0.0057)	0.3368 (0.0057)	0.3322 (0.0058)	0.3424 (0.0060)	0.4389 (0.0074)	0.3628 (0.0110)	0.3320 (0.0058)
T=200	0.2127 (0.0042)	0.2144 (0.0042)	0.2145 (0.0042)	0.2202 (0.0044)	0.3761 (0.0069)	0.2186 (0.0121)	0.2421 (0.0047)
T=500	0.1357 (0.0035)	0.1361 (0.0035)	0.1372 (0.0035)	0.1400 (0.0036)	0.3540 (0.0070)	0.1053 (0.0042)	0.1974 (0.0044)
	$\rho=0.50$						
T=100	0.3272 (0.0054)	0.3347 (0.0054)	0.3291 (0.0055)	0.3377 (0.0057)	0.4368 (0.0074)	0.3948 (0.0194)	0.3292 (0.0057)
T=200	0.2201 (0.0043)	0.2218 (0.0043)	0.2215 (0.0044)	0.2270 (0.0045)	0.3826 (0.0069)	0.2194 (0.0097)	0.2460 (0.0048)
T=500	0.1401 (0.0036)	0.1405 (0.0036)	0.1413 (0.0037)	0.1435 (0.0037)	0.3507 (0.0073)	0.1078 (0.0039)	0.1956 (0.0046)
	$\rho=0.80$						
T=100	0.3529 (0.0058)	0.3627 (0.0059)	0.3547 (0.0059)	0.3628 (0.0060)	0.4585 (0.0076)	0.4161 (0.0110)	0.3476 (0.0058)
T=200	0.2420 (0.0049)	0.2441 (0.0048)	0.2436 (0.0049)	0.2499 (0.0051)	0.3913 (0.0074)	0.2501 (0.0117)	0.2591 (0.0052)
T=500	0.1475 (0.0036)	0.1478 (0.0036)	0.1485 (0.0036)	0.1513 (0.0037)	0.3638 (0.0072)	0.1155 (0.0044)	0.2040 (0.0045)
	$\rho=0.99$						
T=100	1.3642 (0.0514)	1.6216 (0.0605)	1.3155 (0.0508)	1.3468 (0.0515)	0.5319 (0.0076)	1.3814 (0.1203)	0.5141 (0.0129)
T=200	0.7399 (0.0253)	0.8044 (0.0277)	0.7279 (0.0251)	0.7389 (0.0253)	0.4296 (0.0078)	0.7678 (0.1073)	0.3266 (0.0077)
T=500	0.2532 (0.0069)	0.2569 (0.0070)	0.2541 (0.0069)	0.2565 (0.0069)	0.3651 (0.0076)	0.3168 (0.0654)	0.2147 (0.0049)

**Table 4 – Prediction accuracy when coefficients experience one discrete break - Monte Carlo evidence - MSE of one-step-ahead predictions**

**Panel A - One lag in the estimated equation**

	TVC-MA	TVC-MS	TVC- $\Pi$	TVC- $\pi$	OLS	BP	BP-MA
	$\rho=0$						
T=100	1.2975 (0.0272)	1.2989 (0.0275)	1.3030 (0.0277)	1.3243 (0.0286)	1.6029 (0.0349)	1.3322 (0.0566)	1.3784 (0.0278)
T=200	1.1810 (0.0124)	1.1794 (0.0120)	1.1867 (0.0127)	1.1987 (0.0136)	1.5315 (0.0287)	1.1777 (0.0248)	1.2937 (0.0181)
T=500	1.1101 (0.0075)	1.1105 (0.0075)	1.1138 (0.0077)	1.1208 (0.0084)	1.6174 (0.0424)	1.0930 (0.0171)	1.3093 (0.0212)
	$\rho=0.50$						
T=100	1.2645 (0.0144)	1.2643 (0.0147)	1.2753 (0.0157)	1.2991 (0.0168)	1.5421 (0.0247)	1.2941 (0.0240)	1.3343 (0.0161)
T=200	1.2031 (0.0140)	1.2039 (0.0142)	1.2095 (0.0144)	1.2192 (0.0148)	1.6492 (0.0450)	1.1700 (0.0147)	1.3604 (0.0252)
T=500	1.1099 (0.0078)	1.1106 (0.0079)	1.1141 (0.0083)	1.1213 (0.0088)	1.6162 (0.0425)	1.0908 (0.0157)	1.3064 (0.0213)
	$\rho=0.80$						
T=100	1.2814 (0.0143)	1.2815 (0.0141)	1.2882 (0.0146)	1.3202 (0.0168)	1.5781 (0.0242)	1.3859 (0.0367)	1.3552 (0.0167)
T=200	1.1693 (0.0101)	1.1684 (0.0099)	1.1742 (0.0104)	1.1876 (0.0117)	1.5674 (0.0360)	1.1481 (0.0109)	1.2955 (0.0154)
T=500	1.1082 (0.0076)	1.1085 (0.0076)	1.1126 (0.0082)	1.1187 (0.0086)	1.6158 (0.0426)	1.0867 (0.0122)	1.3069 (0.0215)
	$\rho=0.99$						
T=100	1.4014 (0.0235)	1.4357 (0.0241)	1.4051 (0.0236)	1.4245 (0.0241)	1.6442 (0.0416)	1.3933 (0.0480)	1.3764 (0.0277)
T=200	1.2821 (0.0154)	1.2951 (0.0156)	1.2841 (0.0153)	1.2929 (0.0156)	1.5419 (0.0256)	1.2005 (0.0196)	1.3027 (0.0175)
T=500	1.1807 (0.0236)	1.1857 (0.0242)	1.1846 (0.0244)	1.1887 (0.0245)	1.6844 (0.0481)	1.1552 (0.0371)	1.3481 (0.0309)

**Panel B – Three lags in the estimated equation**

	TVC-MA	TVC-MS	TVC- $\Pi$	TVC- $\pi$	OLS	BP	BP-MA
	$\rho=0$						
T=100	1.4007 (0.0171)	1.4022 (0.0161)	1.4042 (0.0172)	1.4272 (0.0215)	1.6117 (0.0263)	1.5408 (0.1302)	1.4496 (0.0217)
T=200	1.2689 (0.0155)	1.2692 (0.0158)	1.2738 (0.0157)	1.2855 (0.0162)	1.5737 (0.0243)	1.3308 (0.0842)	1.3612 (0.0172)
T=500	1.1770 (0.0131)	1.1773 (0.0131)	1.1803 (0.0132)	1.1877 (0.0143)	1.5753 (0.0283)	1.1150 (0.0088)	1.3172 (0.0168)
	$\rho=0.50$						
T=100	1.4022 (0.0184)	1.4059 (0.0183)	1.4070 (0.0186)	1.4255 (0.0196)	1.6490 (0.0282)	1.6550 (0.0968)	1.4823 (0.0231)
T=200	1.2832 (0.0139)	1.2835 (0.0140)	1.2872 (0.0141)	1.2969 (0.0144)	1.6013 (0.0246)	1.3515 (0.0764)	1.3825 (0.0185)
T=500	1.1576 (0.0088)	1.1582 (0.0089)	1.1600 (0.0089)	1.1634 (0.0090)	1.5436 (0.0270)	1.1125 (0.0076)	1.2884 (0.0145)
	$\rho=0.80$						
T=100	1.4206 (0.0222)	1.4245 (0.0224)	1.4242 (0.0223)	1.4426 (0.0227)	1.6930 (0.0325)	1.5160 (0.0584)	1.4954 (0.0231)
T=200	1.2864 (0.0269)	1.2831 (0.0254)	1.2891 (0.0269)	1.3045 (0.0303)	1.6034 (0.0384)	1.2747 (0.0302)	1.3675 (0.0237)
T=500	1.1733 (0.0107)	1.1727 (0.0105)	1.1758 (0.0108)	1.1820 (0.0114)	1.6348 (0.0435)	1.1153 (0.0075)	1.3417 (0.0228)
	$\rho=0.99$						
T=100	1.4718 (0.0199)	1.5054 (0.0206)	1.4852 (0.0204)	1.5134 (0.0212)	1.6644 (0.0332)	1.4521 (0.0274)	1.4802 (0.0234)
T=200	1.3123 (0.0138)	1.3224 (0.0143)	1.3183 (0.0140)	1.3365 (0.0145)	1.6158 (0.0407)	1.3670 (0.1117)	1.3750 (0.0193)
T=500	1.1916 (0.0130)	1.1934 (0.0131)	1.1967 (0.0132)	1.2029 (0.0136)	1.5966 (0.0295)	1.1138 (0.0070)	1.3327 (0.0168)

**Table 5 – Estimation precision when coefficients change every period - Monte Carlo evidence - MSE of coefficient estimates**

**Panel A - One lag in the estimated equation**

	TVC-MA	TVC-MS	TVC-II	TVC- $\pi$	OLS	BP	BP-MA
	$\rho=0$						
T=100	0.1768 (0.0020)	0.1778 (0.0020)	0.1803 (0.0021)	0.1939 (0.0023)	0.3653 (0.0051)	0.2333 (0.0061)	0.2331 (0.0029)
T=200	0.1207 (0.0014)	0.1210 (0.0014)	0.1224 (0.0014)	0.1280 (0.0015)	0.3535 (0.0049)	0.1914 (0.0391)	0.2048 (0.0026)
T=500	0.0718 (0.0008)	0.0718 (0.0008)	0.0722 (0.0008)	0.0735 (0.0008)	0.3409 (0.0048)	0.0925 (0.0015)	0.1766 (0.0022)
	$\rho=0.50$						
T=100	0.1856 (0.0022)	0.1865 (0.0021)	0.1891 (0.0022)	0.2031 (0.0025)	0.3775 (0.0054)	0.2689 (0.0097)	0.2413 (0.0032)
T=200	0.1216 (0.0014)	0.1219 (0.0014)	0.1234 (0.0014)	0.1287 (0.0015)	0.3531 (0.0050)	0.1635 (0.0039)	0.2055 (0.0027)
T=500	0.0719 (0.0008)	0.0720 (0.0008)	0.0724 (0.0008)	0.0735 (0.0008)	0.3490 (0.0050)	0.0982 (0.0020)	0.1770 (0.0023)
	$\rho=0.80$						
T=100	0.2003 (0.0023)	0.2042 (0.0024)	0.2029 (0.0024)	0.2146 (0.0025)	0.3738 (0.0051)	0.3252 (0.0102)	0.2403 (0.0030)
T=200	0.1268 (0.0014)	0.1275 (0.0014)	0.1285 (0.0014)	0.1333 (0.0015)	0.3353 (0.0048)	0.1982 (0.0069)	0.1971 (0.0026)
T=500	0.0733 (0.0008)	0.0733 (0.0008)	0.0737 (0.0008)	0.0747 (0.0008)	0.3413 (0.0048)	0.1108 (0.0034)	0.1759 (0.0023)
	$\rho=0.99$						
T=100	5.8872 (0.1915)	6.5930 (0.2162)	5.8655 (0.1915)	5.8674 (0.1915)	0.4481 (0.0057)	3.8089 (0.3557)	0.4898 (0.0186)
T=200	3.2478 (0.1078)	3.4124 (0.1123)	3.2435 (0.1078)	3.2428 (0.1078)	0.3766 (0.0050)	1.6324 (0.1588)	0.2718 (0.0038)
T=500	0.8609 (0.0329)	0.8916 (0.0340)	0.8608 (0.0329)	0.8609 (0.0329)	0.3395 (0.0047)	0.5805 (0.0686)	0.1835 (0.0023)

**Panel B – Three lags in the estimated equation**

	TVC-MA	TVC-MS	TVC-II	TVC- $\pi$	OLS	BP	BP-MA
	$\rho=0$						
T=100	0.3279 (0.0034)	0.3378 (0.0034)	0.3304 (0.0034)	0.3407 (0.0036)	0.4386 (0.0051)	0.3903 (0.0054)	0.3544 (0.0037)
T=200	0.2187 (0.0021)	0.2221 (0.0021)	0.2202 (0.0022)	0.2247 (0.0022)	0.3810 (0.0049)	0.2514 (0.0032)	0.2723 (0.0030)
T=500	0.1349 (0.0013)	0.1358 (0.0013)	0.1354 (0.0013)	0.1364 (0.0013)	0.3526 (0.0048)	0.1559 (0.0018)	0.2151 (0.0026)
	$\rho=0.50$						
T=100	0.3369 (0.0033)	0.3483 (0.0033)	0.3391 (0.0033)	0.3478 (0.0035)	0.4395 (0.0052)	0.3954 (0.0054)	0.3553 (0.0037)
T=200	0.2279 (0.0022)	0.2313 (0.0022)	0.2292 (0.0022)	0.2339 (0.0023)	0.3834 (0.0049)	0.2607 (0.0034)	0.2753 (0.0031)
T=500	0.1389 (0.0013)	0.1397 (0.0013)	0.1393 (0.0013)	0.1404 (0.0013)	0.3553 (0.0048)	0.1608 (0.0023)	0.2176 (0.0026)
	$\rho=0.80$						
T=100	0.3608 (0.0036)	0.3735 (0.0036)	0.3632 (0.0036)	0.3728 (0.0038)	0.4500 (0.0053)	0.4528 (0.0162)	0.3665 (0.0038)
T=200	0.2397 (0.0023)	0.2442 (0.0023)	0.2410 (0.0023)	0.2456 (0.0024)	0.3901 (0.0050)	0.2756 (0.0040)	0.2806 (0.0031)
T=500	0.1489 (0.0014)	0.1498 (0.0014)	0.1493 (0.0014)	0.1504 (0.0014)	0.3655 (0.0049)	0.1686 (0.0026)	0.2241 (0.0027)
	$\rho=0.99$						
T=100	1.3679 (0.0571)	1.6564 (0.0657)	1.3119 (0.0565)	1.3457 (0.0570)	0.5345 (0.0059)	1.2136 (0.1291)	0.5107 (0.0075)
T=200	0.7702 (0.0284)	0.8373 (0.0307)	0.7648 (0.0284)	0.7690 (0.0284)	0.4231 (0.0051)	0.5511 (0.0441)	0.3365 (0.0052)
T=500	0.2467 (0.0063)	0.2511 (0.0065)	0.2472 (0.0063)	0.2480 (0.0063)	0.3633 (0.0048)	0.2323 (0.0080)	0.2291 (0.0026)



**Table 6 – Prediction accuracy when coefficients change every period - Monte Carlo evidence - MSE of one-step-ahead predictions**

**Panel A - One lag in the estimated equation**

	TVC-MA	TVC-MS	TVC- $\Pi$	TVC- $\pi$	OLS	BP	BP-MA
	$\rho=0$						
T=100	1.2637 (0.0082)	1.2660 (0.0081)	1.2696 (0.0084)	1.2887 (0.0089)	1.5793 (0.0246)	1.3656 (0.0249)	1.3679 (0.0129)
T=200	1.1825 (0.0053)	1.1833 (0.0053)	1.1858 (0.0055)	1.1942 (0.0057)	1.5627 (0.0235)	1.2622 (0.0224)	1.3261 (0.0125)
T=500	1.1094 (0.0032)	1.1096 (0.0032)	1.1101 (0.0032)	1.1119 (0.0033)	1.5546 (0.0240)	1.1515 (0.0085)	1.2784 (0.0114)
	$\rho=0.50$						
T=100	1.2760 (0.0086)	1.2751 (0.0084)	1.2830 (0.0090)	1.3093 (0.0100)	1.5948 (0.0212)	1.4155 (0.0228)	1.3831 (0.0127)
T=200	1.1867 (0.0063)	1.1868 (0.0063)	1.1894 (0.0064)	1.2000 (0.0071)	1.5554 (0.0212)	1.2433 (0.0098)	1.3298 (0.0124)
T=500	1.1161 (0.0032)	1.1159 (0.0032)	1.1170 (0.0033)	1.1192 (0.0034)	1.5943 (0.0245)	1.1564 (0.0056)	1.3054 (0.0125)
	$\rho=0.80$						
T=100	1.3009 (0.0135)	1.3026 (0.0130)	1.3070 (0.0137)	1.3339 (0.0149)	1.6125 (0.0344)	1.4257 (0.0170)	1.4093 (0.0215)
T=200	1.1894 (0.0128)	1.1903 (0.0131)	1.1926 (0.0130)	1.2047 (0.0141)	1.5493 (0.0251)	1.2723 (0.0180)	1.3286 (0.0161)
T=500	1.1116 (0.0035)	1.1115 (0.0034)	1.1123 (0.0035)	1.1139 (0.0036)	1.5399 (0.0186)	1.1672 (0.0114)	1.2843 (0.0098)
	$\rho=0.99$						
T=100	1.4028 (0.0115)	1.4363 (0.0124)	1.4067 (0.0118)	1.4220 (0.0120)	1.6146 (0.0208)	1.4992 (0.0517)	1.3924 (0.0123)
T=200	1.2848 (0.0091)	1.2972 (0.0093)	1.2867 (0.0091)	1.2917 (0.0092)	1.5963 (0.0224)	1.2675 (0.0117)	1.3379 (0.0122)
T=500	1.1564 (0.0044)	1.1596 (0.0045)	1.1574 (0.0045)	1.1588 (0.0045)	1.5581 (0.0231)	1.1752 (0.0151)	1.2882 (0.0115)

**Panel B – Three lags in the estimated equation**

	TVC-MA	TVC-MS	TVC- $\Pi$	TVC- $\pi$	OLS	BP	BP-MA
	$\rho=0$						
T=100	1.4312 (0.0188)	1.4333 (0.0168)	1.4359 (0.0191)	1.4636 (0.0230)	1.6667 (0.0306)	1.5028 (0.0241)	1.5340 (0.0256)
T=200	1.2653 (0.0071)	1.2653 (0.0068)	1.2681 (0.0072)	1.2800 (0.0082)	1.5577 (0.0173)	1.3186 (0.0107)	1.3912 (0.0116)
T=500	1.1713 (0.0051)	1.1715 (0.0051)	1.1723 (0.0052)	1.1737 (0.0052)	1.5745 (0.0226)	1.2156 (0.0077)	1.3453 (0.0135)
	$\rho=0.50$						
T=100	1.4214 (0.0137)	1.4247 (0.0131)	1.4252 (0.0138)	1.4476 (0.0152)	1.6373 (0.0229)	1.4982 (0.0197)	1.5109 (0.0189)
T=200	1.2954 (0.0128)	1.2961 (0.0127)	1.2993 (0.0136)	1.3114 (0.0141)	1.6280 (0.0247)	1.3697 (0.0214)	1.4369 (0.0175)
T=500	1.1687 (0.0048)	1.1689 (0.0048)	1.1695 (0.0048)	1.1711 (0.0049)	1.5631 (0.0197)	1.2145 (0.0071)	1.3402 (0.0115)
	$\rho=0.80$						
T=100	1.4303 (0.0129)	1.4342 (0.0125)	1.4353 (0.0131)	1.4614 (0.0146)	1.6616 (0.0246)	1.5534 (0.0261)	1.5281 (0.0187)
T=200	1.2803 (0.0079)	1.2816 (0.0078)	1.2829 (0.0079)	1.2941 (0.0083)	1.5949 (0.0208)	1.3301 (0.0099)	1.4126 (0.0131)
T=500	1.1652 (0.0044)	1.1648 (0.0043)	1.1659 (0.0044)	1.1681 (0.0044)	1.5584 (0.0233)	1.2240 (0.0120)	1.3402 (0.0134)
	$\rho=0.99$						
T=100	1.5138 (0.0153)	1.5417 (0.0162)	1.5246 (0.0156)	1.5664 (0.0167)	1.6848 (0.0276)	1.5574 (0.0300)	1.5380 (0.0200)
T=200	1.3405 (0.0110)	1.3533 (0.0113)	1.3469 (0.0112)	1.3624 (0.0116)	1.6228 (0.0267)	1.3507 (0.0122)	1.4279 (0.0152)
T=500	1.1838 (0.0071)	1.1851 (0.0071)	1.1853 (0.0071)	1.1889 (0.0072)	1.5842 (0.0276)	1.2445 (0.0173)	1.3475 (0.0109)

**Table 7 – Prediction accuracy – Risk exposures – Reduction in the MSE of one-step-ahead predictions (benchmark=OLS)**

	TVC-MA	TVC-MS	TVC- <i>II</i>	TVC- $\pi$	OLS	BP	BP-MA
Mean	2.94%	2.48%	2.70%	2.15%	0%	-121.72%	3.13%
Standard dev.	11.02%	11.68%	10.82%	10.63%	0%	-557.07%	6.00%
First quartile	-2.14%	-2.67%	-2.14%	-2.23%	0%	-46.94%	-0.14%
Median	1.85%	1.30%	1.17%	0.11%	0%	-5.76%	1.53%
Third quartile	7.83%	7.75%	7.47%	6.84%	0%	0.20%	6.37%

RECENTLY PUBLISHED “TEMI” (\*)

- N. 811 – *Schooling and youth mortality: learning from a mass military exemption*, by Piero Cipollone and Alfonso Rosolia (June 2011).
- N. 812 – *Welfare costs of inflation and the circulation of US currency abroad*, by Alessandro Calza and Andrea Zaghini (June 2011).
- N. 813 – *Legal status of immigrants and criminal behavior: evidence from a natural experiment*, by Giovanni Mastrobuoni and Paolo Pinotti (June 2011).
- N. 814 – *An unexpected crisis? Looking at pricing effectiveness of different banks*, by Valerio Vacca (July 2011).
- N. 815 – *Skills or culture? An analysis of the decision to work by immigrant women in Italy*, by Antonio Accetturo and Luigi Infante (July 2011).
- N. 816 – *Home bias in interbank lending and banks’ resolution regimes*, by Michele Manna (July 2011).
- N. 817 – *Macroeconomic determinants of carry trade activity*, by Alessio Anzuini and Fabio Fornari (September 2011).
- N. 818 – *Leaving home and housing prices. The experience of Italian youth emancipation*, by Francesca Modena and Concetta Rondinelli (September 2011).
- N. 819 – *The interbank market after the financial turmoil: squeezing liquidity in a “lemons market” or asking liquidity “on tap”*, by Antonio De Socio (September 2011).
- N. 820 – *The relationship between the PMI and the Italian index of industrial production and the impact of the latest economic crisis*, by Valentina Aprigliano (September 2011).
- N. 821 – *Inside the sovereign credit default swap market: price discovery, announcements, market behaviour and corporate sector*, by Alessandro Carboni (September 2011).
- N. 822 – *The demand for energy of Italian households*, by Ivan Faiella (September 2011).
- N. 823 – *Sull’ampiezza ottimale delle giurisdizioni locali: il caso delle province italiane*, by Guglielmo Barone (September 2011).
- N. 824 – *The public-private pay gap: a robust quantile approach*, by Domenico Depalo and Raffaella Giordano (September 2011).
- N. 825 – *Evaluating students’ evaluations of professors*, by Michele Braga, Marco Paccagnella and Michele Pellizzari (October 2011).
- N. 826 – *Do interbank customer relationships exist? And how did they function in the crisis? Learning from Italy*, by Massimiliano Affinito (October 2011).
- N. 827 – *Foreign trade, home linkages and the spatial transmission of economic fluctuations in Italy*, by Valter Di Giacinto (October 2011).
- N. 828 – *Healthcare in Italy: expenditure determinants and regional differentials*, by Maura Francese and Marzia Romanelli (October 2011).
- N. 829 – *Bank heterogeneity and interest rate setting: what lessons have we learned since Lehman Brothers?*, by Leonardo Gambacorta and Paolo Emilio Mistrulli (October 2011).
- N. 830 – *Structural reforms and macroeconomic performance in the euro area countries: a model-based assessment*, by Sandra Gomes, Pascal Jacquinot, Matthias Mohr and Massimiliano Pisani (October 2011).
- N. 831 – *Risk measures for autocorrelated hedge fund returns*, by Antonio Di Cesare, Philip A. Stork and Casper G. de Vries (October 2011).

---

(\*) Requests for copies should be sent to:  
Banca d’Italia – Servizio Studi di struttura economica e finanziaria – Divisione Biblioteca e Archivio storico – Via Nazionale, 91 – 00184 Rome – (fax 0039 06 47922059). They are available on the Internet [www.bancaditalia.it](http://www.bancaditalia.it).

2008

- P. ANGELINI, *Liquidity and announcement effects in the euro area*, *Giornale degli Economisti e Annali di Economia*, v. 67, 1, pp. 1-20, **TD No. 451 (October 2002)**.
- P. ANGELINI, P. DEL GIOVANE, S. SIVIERO and D. TERLIZZESE, *Monetary policy in a monetary union: What role for regional information?*, *International Journal of Central Banking*, v. 4, 3, pp. 1-28, **TD No. 457 (December 2002)**.
- F. SCHIVARDI and R. TORRINI, *Identifying the effects of firing restrictions through size-contingent Differences in regulation*, *Labour Economics*, v. 15, 3, pp. 482-511, **TD No. 504 (June 2004)**.
- L. GUIISO and M. PAIELLA., *Risk aversion, wealth and background risk*, *Journal of the European Economic Association*, v. 6, 6, pp. 1109-1150, **TD No. 483 (September 2003)**.
- C. BIANCOTTI, G. D'ALESSIO and A. NERI, *Measurement errors in the Bank of Italy's survey of household income and wealth*, *Review of Income and Wealth*, v. 54, 3, pp. 466-493, **TD No. 520 (October 2004)**.
- S. MOMIGLIANO, J. HENRY and P. HERNÁNDEZ DE COS, *The impact of government budget on prices: Evidence from macroeconomic models*, *Journal of Policy Modelling*, v. 30, 1, pp. 123-143 **TD No. 523 (October 2004)**.
- L. GAMBACORTA, *How do banks set interest rates?*, *European Economic Review*, v. 52, 5, pp. 792-819, **TD No. 542 (February 2005)**.
- P. ANGELINI and A. GENERALE, *On the evolution of firm size distributions*, *American Economic Review*, v. 98, 1, pp. 426-438, **TD No. 549 (June 2005)**.
- R. FELICI and M. PAGNINI, *Distance, bank heterogeneity and entry in local banking markets*, *The Journal of Industrial Economics*, v. 56, 3, pp. 500-534, **No. 557 (June 2005)**.
- S. DI ADDARIO and E. PATACCHINI, *Wages and the city. Evidence from Italy*, *Labour Economics*, v.15, 5, pp. 1040-1061, **TD No. 570 (January 2006)**.
- S. SCALIA, *Is foreign exchange intervention effective?*, *Journal of International Money and Finance*, v. 27, 4, pp. 529-546, **TD No. 579 (February 2006)**.
- M. PERICOLI and M. TABOGA, *Canonical term-structure models with observable factors and the dynamics of bond risk premia*, *Journal of Money, Credit and Banking*, v. 40, 7, pp. 1471-88, **TD No. 580 (February 2006)**.
- E. VIVIANO, *Entry regulations and labour market outcomes. Evidence from the Italian retail trade sector*, *Labour Economics*, v. 15, 6, pp. 1200-1222, **TD No. 594 (May 2006)**.
- S. FEDERICO and G. A. MINERVA, *Outward FDI and local employment growth in Italy*, *Review of World Economics*, *Journal of Money, Credit and Banking*, v. 144, 2, pp. 295-324, **TD No. 613 (February 2007)**.
- F. Busetti and A. HARVEY, *Testing for trend*, *Econometric Theory*, v. 24, 1, pp. 72-87, **TD No. 614 (February 2007)**.
- V. CESTARI, P. DEL GIOVANE and C. ROSSI-ARNAUD, *Memory for prices and the Euro cash changeover: an analysis for cinema prices in Italy*, In P. Del Giovane e R. Sabbatini (eds.), *The Euro Inflation and Consumers' Perceptions. Lessons from Italy*, Berlin-Heidelberg, Springer, **TD No. 619 (February 2007)**.
- B. H. HALL, F. LOTTI and J. MAIRESSE, *Employment, innovation and productivity: evidence from Italian manufacturing microdata*, *Industrial and Corporate Change*, v. 17, 4, pp. 813-839, **TD No. 622 (April 2007)**.
- J. SOUSA and A. ZAGHINI, *Monetary policy shocks in the Euro Area and global liquidity spillovers*, *International Journal of Finance and Economics*, v.13, 3, pp. 205-218, **TD No. 629 (June 2007)**.
- M. DEL GATTO, GIANMARCO I. P. OTTAVIANO and M. PAGNINI, *Openness to trade and industry cost dispersion: Evidence from a panel of Italian firms*, *Journal of Regional Science*, v. 48, 1, pp. 97-129, **TD No. 635 (June 2007)**.
- P. DEL GIOVANE, S. FABIANI and R. SABBATINI, *What's behind "inflation perceptions"? A survey-based analysis of Italian consumers*, in P. Del Giovane e R. Sabbatini (eds.), *The Euro Inflation and Consumers' Perceptions. Lessons from Italy*, Berlin-Heidelberg, Springer, **TD No. 655 (January 2008)**.
- R. BRONZINI, G. DE BLASIO, G. PELLEGRINI and A. SCOGNAMIGLIO, *La valutazione del credito d'imposta per gli investimenti*, *Rivista di politica economica*, v. 98, 4, pp. 79-112, **TD No. 661 (April 2008)**.

- B. BORTOLOTTI, and P. PINOTTI, *Delayed privatization*, Public Choice, v. 136, 3-4, pp. 331-351, **TD No. 663 (April 2008)**.
- R. BONCI and F. COLUMBA, *Monetary policy effects: New evidence from the Italian flow of funds*, Applied Economics, v. 40, 21, pp. 2803-2818, **TD No. 678 (June 2008)**.
- M. CUCCULELLI, and G. MICUCCI, *Family Succession and firm performance: evidence from Italian family firms*, Journal of Corporate Finance, v. 14, 1, pp. 17-31, **TD No. 680 (June 2008)**.
- A. SILVESTRINI and D. VEREDAS, *Temporal aggregation of univariate and multivariate time series models: a survey*, Journal of Economic Surveys, v. 22, 3, pp. 458-497, **TD No. 685 (August 2008)**.

2009

- F. PANETTA, F. SCHIVARDI and M. SHUM, *Do mergers improve information? Evidence from the loan market*, Journal of Money, Credit, and Banking, v. 41, 4, pp. 673-709, **TD No. 521 (October 2004)**.
- M. BUGAMELLI and F. PATERNÒ, *Do workers' remittances reduce the probability of current account reversals?*, World Development, v. 37, 12, pp. 1821-1838, **TD No. 573 (January 2006)**.
- P. PAGANO and M. PISANI, *Risk-adjusted forecasts of oil prices*, The B.E. Journal of Macroeconomics, v. 9, 1, Article 24, **TD No. 585 (March 2006)**.
- M. PERICOLI and M. SBRACIA, *The CAPM and the risk appetite index: theoretical differences, empirical similarities, and implementation problems*, International Finance, v. 12, 2, pp. 123-150, **TD No. 586 (March 2006)**.
- U. ALBERTAZZI and L. GAMBACORTA, *Bank profitability and the business cycle*, Journal of Financial Stability, v. 5, 4, pp. 393-409, **TD No. 601 (September 2006)**.
- S. MAGRI, *The financing of small innovative firms: the Italian case*, Economics of Innovation and New Technology, v. 18, 2, pp. 181-204, **TD No. 640 (September 2007)**.
- V. DI GIACINTO and G. MICUCCI, *The producer service sector in Italy: long-term growth and its local determinants*, Spatial Economic Analysis, Vol. 4, No. 4, pp. 391-425, **TD No. 643 (September 2007)**.
- F. LORENZO, L. MONTEFORTE and L. SESSA, *The general equilibrium effects of fiscal policy: estimates for the euro area*, Journal of Public Economics, v. 93, 3-4, pp. 559-585, **TD No. 652 (November 2007)**.
- R. GOLINELLI and S. MOMIGLIANO, *The Cyclical Reaction of Fiscal Policies in the Euro Area. A Critical Survey of Empirical Research*, Fiscal Studies, v. 30, 1, pp. 39-72, **TD No. 654 (January 2008)**.
- P. DEL GIOVANE, S. FABIANI and R. SABBATINI, *What's behind "Inflation Perceptions"? A survey-based analysis of Italian consumers*, Giornale degli Economisti e Annali di Economia, v. 68, 1, pp. 25-52, **TD No. 655 (January 2008)**.
- F. MACCHERONI, M. MARINACCI, A. RUSTICHINI and M. TABOGA, *Portfolio selection with monotone mean-variance preferences*, Mathematical Finance, v. 19, 3, pp. 487-521, **TD No. 664 (April 2008)**.
- M. AFFINITO and M. PIAZZA, *What are borders made of? An analysis of barriers to European banking integration*, in P. Alessandrini, M. Fratianni and A. Zazzaro (eds.): The Changing Geography of Banking and Finance, Dordrecht Heidelberg London New York, Springer, **TD No. 666 (April 2008)**.
- A. BRANDOLINI, *On applying synthetic indices of multidimensional well-being: health and income inequalities in France, Germany, Italy, and the United Kingdom*, in R. Gotoh and P. Dumouchel (eds.), Against Injustice. The New Economics of Amartya Sen, Cambridge, Cambridge University Press, **TD No. 668 (April 2008)**.
- G. FERRERO and A. NOBILI, *Futures contract rates as monetary policy forecasts*, International Journal of Central Banking, v. 5, 2, pp. 109-145, **TD No. 681 (June 2008)**.
- P. CASADIO, M. LO CONTE and A. NERI, *Balancing work and family in Italy: the new mothers' employment decisions around childbearing*, in T. Addabbo and G. Solinas (eds.), Non-Standard Employment and Quality of Work, Physica-Verlag. A Springer Company, **TD No. 684 (August 2008)**.
- L. ARCIERO, C. BIANCOTTI, L. D'AURIZIO and C. IMPENNA, *Exploring agent-based methods for the analysis of payment systems: A crisis model for StarLogo TNG*, Journal of Artificial Societies and Social Simulation, v. 12, 1, **TD No. 686 (August 2008)**.
- A. CALZA and A. ZAGHINI, *Nonlinearities in the dynamics of the euro area demand for M1*, Macroeconomic Dynamics, v. 13, 1, pp. 1-19, **TD No. 690 (September 2008)**.
- L. FRANCESCO and A. SECCHI, *Technological change and the households' demand for currency*, Journal of Monetary Economics, v. 56, 2, pp. 222-230, **TD No. 697 (December 2008)**.
- G. ASCARI and T. ROPELE, *Trend inflation, taylor principle, and indeterminacy*, Journal of Money, Credit and Banking, v. 41, 8, pp. 1557-1584, **TD No. 708 (May 2007)**.

- S. COLAROSSO and A. ZAGHINI, *Gradualism, transparency and the improved operational framework: a look at overnight volatility transmission*, *International Finance*, v. 12, 2, pp. 151-170, **TD No. 710 (May 2009)**.
- M. BUGAMELLI, F. SCHIVARDI and R. ZIZZA, *The euro and firm restructuring*, in A. Alesina e F. Giavazzi (eds): *Europe and the Euro*, Chicago, University of Chicago Press, **TD No. 716 (June 2009)**.
- B. HALL, F. LOTTI and J. MAIRESSE, *Innovation and productivity in SMEs: empirical evidence for Italy*, *Small Business Economics*, v. 33, 1, pp. 13-33, **TD No. 718 (June 2009)**.

2010

- A. PRATI and M. SBRACIA, *Uncertainty and currency crises: evidence from survey data*, *Journal of Monetary Economics*, v. 57, 6, pp. 668-681, **TD No. 446 (July 2002)**.
- L. MONTEFORTE and S. SIVIERO, *The Economic Consequences of Euro Area Modelling Shortcuts*, *Applied Economics*, v. 42, 19-21, pp. 2399-2415, **TD No. 458 (December 2002)**.
- S. MAGRI, *Debt maturity choice of nonpublic Italian firms*, *Journal of Money, Credit, and Banking*, v.42, 2-3, pp. 443-463, **TD No. 574 (January 2006)**.
- R. BRONZINI and P. PISELLI, *Determinants of long-run regional productivity with geographical spillovers: the role of R&D, human capital and public infrastructure*, *Regional Science and Urban Economics*, v. 39, 2, pp.187-199, **TD No. 597 (September 2006)**.
- E. IOSSA and G. PALUMBO, *Over-optimism and lender liability in the consumer credit market*, *Oxford Economic Papers*, v. 62, 2, pp. 374-394, **TD No. 598 (September 2006)**.
- S. NERI and A. NOBILI, *The transmission of US monetary policy to the euro area*, *International Finance*, v. 13, 1, pp. 55-78, **TD No. 606 (December 2006)**.
- F. ALTISSIMO, R. CRISTADORO, M. FORNI, M. LIPPI and G. VERONESE, *New Eurocoin: Tracking Economic Growth in Real Time*, *Review of Economics and Statistics*, v. 92, 4, pp. 1024-1034, **TD No. 631 (June 2007)**.
- A. CIARLONE, P. PISELLI and G. TREBESCHI, *Emerging Markets' Spreads and Global Financial Conditions*, *Journal of International Financial Markets, Institutions & Money*, v. 19, 2, pp. 222-239, **TD No. 637 (June 2007)**.
- U. ALBERTAZZI and L. GAMBACORTA, *Bank profitability and taxation*, *Journal of Banking and Finance*, v. 34, 11, pp. 2801-2810, **TD No. 649 (November 2007)**.
- M. IACOVIELLO and S. NERI, *Housing market spillovers: evidence from an estimated DSGE model*, *American Economic Journal: Macroeconomics*, v. 2, 2, pp. 125-164, **TD No. 659 (January 2008)**.
- F. BALASSONE, F. MAURA and S. ZOTTERI, *Cyclical asymmetry in fiscal variables in the EU*, *Empirica*, **TD No. 671**, v. 37, 4, pp. 381-402 **(June 2008)**.
- F. D'AMURI, O. GIANMARCO I.P. and P. GIOVANNI, *The labor market impact of immigration on the western german labor market in the 1990s*, *European Economic Review*, v. 54, 4, pp. 550-570, **TD No. 687 (August 2008)**.
- A. ACCETTURO, *Agglomeration and growth: the effects of commuting costs*, *Papers in Regional Science*, v. 89, 1, pp. 173-190, **TD No. 688 (September 2008)**.
- S. NOBILI and G. PALAZZO, *Explaining and forecasting bond risk premiums*, *Financial Analysts Journal*, v. 66, 4, pp. 67-82, **TD No. 689 (September 2008)**.
- A. B. ATKINSON and A. BRANDOLINI, *On analysing the world distribution of income*, *World Bank Economic Review*, v. 24, 1, pp. 1-37, **TD No. 701 (January 2009)**.
- R. CAPPARIELLO and R. ZIZZA, *Dropping the Books and Working Off the Books*, *Labour*, v. 24, 2, pp. 139-162, **TD No. 702 (January 2009)**.
- C. NICOLETTI and C. RONDINELLI, *The (mis)specification of discrete duration models with unobserved heterogeneity: a Monte Carlo study*, *Journal of Econometrics*, v. 159, 1, pp. 1-13, **TD No. 705 (March 2009)**.
- L. FORNI, A. GERALI and M. PISANI, *Macroeconomic effects of greater competition in the service sector: the case of Italy*, *Macroeconomic Dynamics*, v. 14, 5, pp. 677-708, **TD No. 706 (March 2009)**.
- V. DI GIACINTO, G. MICUCCI and P. MONTANARO, *Dynamic macroeconomic effects of public capital: evidence from regional Italian data*, *Giornale degli economisti e annali di economia*, v. 69, 1, pp. 29-66, **TD No. 733 (November 2009)**.
- F. COLUMBA, L. GAMBACORTA and P. E. MISTRULLI, *Mutual Guarantee institutions and small business finance*, *Journal of Financial Stability*, v. 6, 1, pp. 45-54, **TD No. 735 (November 2009)**.

- A. GERALI, S. NERI, L. SESSA and F. M. SIGNORETTI, *Credit and banking in a DSGE model of the Euro Area*, Journal of Money, Credit and Banking, v. 42, 6, pp. 107-141, **TD No. 740 (January 2010)**.
- M. AFFINITO and E. TAGLIAFERRI, *Why do (or did?) banks securitize their loans? Evidence from Italy*, Journal of Financial Stability, v. 6, 4, pp. 189-202, **TD No. 741 (January 2010)**.
- S. FEDERICO, *Outsourcing versus integration at home or abroad and firm heterogeneity*, Empirica, v. 37, 1, pp. 47-63, **TD No. 742 (February 2010)**.
- V. DI GIACINTO, *On vector autoregressive modeling in space and time*, Journal of Geographical Systems, v. 12, 2, pp. 125-154, **TD No. 746 (February 2010)**.
- S. MOCETTI and C. PORELLO, *How does immigration affect native internal mobility? new evidence from Italy*, Regional Science and Urban Economics, v. 40, 6, pp. 427-439, **TD No. 748 (March 2010)**.
- A. DI CESARE and G. GUAZZAROTTI, *An analysis of the determinants of credit default swap spread changes before and during the subprime financial turmoil*, Journal of Current Issues in Finance, Business and Economics, v. 3, 4, pp., **TD No. 749 (March 2010)**.
- P. CIPOLLONE, P. MONTANARO and P. SESTITO, *Value-added measures in Italian high schools: problems and findings*, Giornale degli economisti e annali di economia, v. 69, 2, pp. 81-114, **TD No. 754 (March 2010)**.
- A. BRANDOLINI, S. MAGRI and T. M. SMEEDING, *Asset-based measurement of poverty*, Journal of Policy Analysis and Management, v. 29, 2, pp. 267-284, **TD No. 755 (March 2010)**.
- G. CAPPELLETTI, *A Note on rationalizability and restrictions on beliefs*, The B.E. Journal of Theoretical Economics, v. 10, 1, pp. 1-11, **TD No. 757 (April 2010)**.
- S. DI ADDARIO and D. VURI, *Entrepreneurship and market size. the case of young college graduates in Italy*, Labour Economics, v. 17, 5, pp. 848-858, **TD No. 775 (September 2010)**.
- A. CALZA and A. ZAGHINI, *Sectoral money demand and the great disinflation in the US*, Journal of Money, Credit, and Banking, v. 42, 8, pp. 1663-1678, **TD No. 785 (January 2011)**.

2011

- S. DI ADDARIO, *Job search in thick markets*, Journal of Urban Economics, v. 69, 3, pp. 303-318, **TD No. 605 (December 2006)**.
- E. CIAPANNA, *Directed matching with endogenous markov probability: clients or competitors?*, The RAND Journal of Economics, v. 42, 1, pp. 92-120, **TD No. 665 (April 2008)**.
- L. FORNI, A. GERALI and M. PISANI, *The Macroeconomics of Fiscal Consolidation in a Monetary Union: the Case of Italy*, in Luigi Paganetto (ed.), Recovery after the crisis. Perspectives and policies, VDM Verlag Dr. Muller, **TD No. 747 (March 2010)**.
- A. DI CESARE and G. GUAZZAROTTI, *An analysis of the determinants of credit default swap changes before and during the subprime financial turmoil*, in Barbara L. Campos and Janet P. Wilkins (eds.), The Financial Crisis: Issues in Business, Finance and Global Economics, New York, Nova Science Publishers, Inc., **TD No. 749 (March 2010)**.
- G. GRANDE and I. VISCO, *A public guarantee of a minimum return to defined contribution pension scheme members*, The Journal of Risk, v. 13, 3, pp. 3-43, **TD No. 762 (June 2010)**.
- P. DEL GIOVANE, G. ERAMO and A. NOBILI, *Disentangling demand and supply in credit developments: a survey-based analysis for Italy*, Journal of Banking and Finance, v. 35, 10, pp. 2719-2732, **TD No. 764 (June 2010)**.
- M. TABOGA, *Under/over-valuation of the stock market and cyclically adjusted earnings*, International Finance, v. 14, 1, pp. 135-164, **TD No. 780 (December 2010)**.
- S. NERI, *Housing, consumption and monetary policy: how different are the U.S. and the Euro area?*, Journal of Banking and Finance, v.35, 11, pp. 3019-3041, **TD No. 807 (April 2011)**.

FORTHCOMING

- M. BUGAMELLI and A. ROSOLIA, *Produttività e concorrenza estera*, Rivista di politica economica, **TD No. 578 (February 2006)**.
- G. DE BLASIO and G. NUZZO, *Historical traditions of civiness and local economic development*, Journal of Regional Science, **TD No. 591 (May 2006)**.

- F. CINGANO and A. ROSOLIA, *People I know: job search and social networks*, Journal of Labor Economics, **TD No. 600 (September 2006)**.
- F. SCHIVARDI and E. VIVIANO, *Entry barriers in retail trade*, Economic Journal, **TD No. 616 (February 2007)**.
- G. FERRERO, A. NOBILI and P. PASSIGLIA, *Assessing excess liquidity in the Euro Area: the role of sectoral distribution of money*, Applied Economics, **TD No. 627 (April 2007)**.
- P. E. MISTRULLI, *Assessing financial contagion in the interbank market: maximum entropy versus observed interbank lending patterns*, Journal of Banking & Finance, **TD No. 641 (September 2007)**.
- Y. ALTUNBAS, L. GAMBACORTA and D. MARQUÉS, *Securitisation and the bank lending channel*, European Economic Review, **TD No. 653 (November 2007)**.
- M. BUGAMELLI and F. PATERNÒ, *Output growth volatility and remittances*, Economica, **TD No. 673 (June 2008)**.
- V. DI GIACINTO e M. PAGNINI, *Local and global agglomeration patterns: two econometrics-based indicators*, Regional Science and Urban Economics, **TD No. 674 (June 2008)**.
- G. BARONE and F. CINGANO, *Service regulation and growth: evidence from OECD countries*, Economic Journal, **TD No. 675 (June 2008)**.
- S. MOCETTI, *Educational choices and the selection process before and after compulsory school*, Education Economics, **TD No. 691 (September 2008)**.
- P. SESTITO and E. VIVIANO, *Reservation wages: explaining some puzzling regional patterns*, Labour, **TD No. 696 (December 2008)**.
- P. PINOTTI, M. BIANCHI and P. BUONANNO, *Do immigrants cause crime?*, Journal of the European Economic Association, **TD No. 698 (December 2008)**.
- R. GIORDANO and P. TOMMASINO, *What determines debt intolerance? The role of political and monetary institutions*, European Journal of Political Economy, **TD No. 700 (January 2009)**.
- F. LIPPI and A. NOBILI, *Oil and the macroeconomy: a quantitative structural analysis*, Journal of European Economic Association, **TD No. 704 (March 2009)**.
- F. CINGANO and P. PINOTTI, *Politicians at work. The private returns and social costs of political connections*, Journal of the European Economic Association, **TD No. 709 (May 2009)**.
- Y. ALTUNBAS, L. GAMBACORTA, and D. MARQUÉS-IBÁÑEZ, *Bank risk and monetary policy*, Journal of Financial Stability, **TD No. 712 (May 2009)**.
- P. ANGELINI, A. NOBILI e C. PICILLO, *The interbank market after August 2007: What has changed, and why?*, Journal of Money, Credit and Banking, **TD No. 731 (October 2009)**.
- G. BARONE and S. MOCETTI, *Tax morale and public spending inefficiency*, International Tax and Public Finance, **TD No. 732 (November 2009)**.
- L. FORNI, A. GERALI and M. PISANI, *The macroeconomics of fiscal consolidations in euro area countries*, Journal of Economic Dynamics and Control, **TD No. 747 (March 2010)**.
- G. BARONE, R. FELICI and M. PAGNINI, *Switching costs in local credit markets*, International Journal of Industrial Organization, **TD No. 760 (June 2010)**.
- G. BARONE and S. MOCETTI, *With a little help from abroad: the effect of low-skilled immigration on the female labour supply*, Labour Economics, **TD No. 766 (July 2010)**.
- S. MAGRI and R. PICO, *The rise of risk-based pricing of mortgage interest rates in Italy*, Journal of Banking and Finance, **TD No. 778 (October 2010)**.
- A. ACCETTURO and G. DE BLASIO, *Policies for local development: an evaluation of Italy's "Patti Territoriali"*, Regional Science and Urban Economics, **TD No. 789 (January 2006)**.
- E. COCOZZA and P. PISELLI, *Testing for east-west contagion in the European banking sector during the financial crisis*, in R. Matoušek; D. Stavárek (eds.), Financial Integration in the European Union, Taylor & Francis, **TD No. 790 (February 2011)**.
- S. NERI and T. ROPELE, *Imperfect information, real-time data and monetary policy in the Euro area*, The Economic Journal, **TD No. 802 (March 2011)**.