Cross-Asset Contagion in the Financial Crisis: A Bayesian Time-Varying Parameter Approach

Massimo Guidolin* Erwin Hansen† Manuela Pedio‡

December 12, 2017

Abstract

We study cross-asset contagion mechanisms in the US financial markets. The recent US subprime crisis provides us with one exogenous shock in a specific market (mortgage-backed securities) to measure contagion. We model the dynamic linkages among markets and allow for changes in this relationship to capture contagion. We look at how and to what extent a negative shock that initially occurred in the asset-backed security (ABS) low-quality market propagated to ABS higher grade, Treasury repos, Treasury note, corporate bond, and stock markets. We rely on dynamic time series models estimated with Bayesian methods. We estimate several specifications ranging from single-state vector autoregressive (VAR) models with constant parameters to fully flexible VAR models where the parameters may vary at each observation. We provide evidence of structural changes in the cross-asset relationships and therefore of contagion. Moreover, by observing the impulse response functions of the models, we conclude that contagion mainly occurred through the flight-to-liquidity, risk premium, and correlated information channels.

JEL Classification: E40, E52.

Key Words: Contagion, bond yield, financial crisis, interdependence, Bayesian estimation.

1 Introduction

The instability in macroeconomic and financial time series as well as in their linkages has been extensively studied in the literature (see, e.g., Sensier and van Dijck, 2004; Stock and Watson, 1996). In particular, when applied to financial time series, such instabilities have inspired a vast empirical literature on contagion, a state of alterations of such linkages that would be characterized by their strengthening in excess to either some notion of "normality" or of fundamental connections (see, e.g., Ciccarelli and Rebucci, 2003; Dornbusch, Park, and Claessens, 2000; Pericoli and Sbracia, 2003; Jawadi, Louhichi, and Cheffou, 2015). However, most of the literature has focused on cross-country relationships, in particular, on episodes of international contagion (see, e.g., Dungey, Fry, Gonzalez-Hermosillo, and

*Corresponding author: Department of Finance, BAFFI-CAREFIN, and IGIER, Bocconi University, Milan, Italy. Email: massimo.guidolin@unibocconi.it.

†Facultad de Economía y Negocios, Universidad de Chile. Address: Diagonal Paraguay 257, Oficina 1204, Santiago, Chile. Tel: +56(2)29772125. Email:ehansen@fen.uchile.cl

‡Department of Finance, Bocconi University, Milan, Italy. Email: manuela.pedio@unibocconi.it.

Martin, 2007; King and Wadhwani, 1990; Kodres and Pritsker, 2002; Reinhart, Kaminsky and Vegh, 2003). In this paper, we perform an empirical investigation of the instability of financial linkages and of contagion effects, but we focus instead on cross-asset, within-country, contagion mechanisms.[1] We exploit the recent US subprime crisis to identify how and to what extent an exogenous shock to (a segment of) the mortgage-backed securities market, may have spilled over to other fixed-income markets and the stock market.[2] As noted by Pesaran and Pick (2007) among others, the distinction between contagion and interdependence has important implications both for investors, because identifying one phenomenon with sensible precision would allow them to better adjust their portfolio diversification strategies, and for policy-makers, as the efficacy of policy interventions highly depends on the nature of the shocks transmission channels. We propose a novel empirical strategy to test whether the propagation effects from asset-backed securities (henceforth, ABS) to other fixed income markets might have been anticipated because "normal", or whether to the contrary such spillovers were "abnormal", hence corresponding to breaks in established relationships and possibly to a cross-market, subprime crisis-induced contagion episode.

The literature provides a variety of different definitions and measures of contagion. Forbes and Rigobon (2002) define contagion as a significant change in cross-market linkages. Accordingly, a strong association between two markets does not by itself give evidence of contagion, as it is only the result of the high level of interdependence that has always characterized these markets. We follow Ciccarelli and Rebucci (2003) who narrow the scope of this definition and define contagion as a temporary change in cross-markets linkages in order to distinguish temporary shifts in cross-markets relations that may occur during a crisis from permanent changes in the transmission mechanisms, namely "structural breaks" (also see the discussion in Abbate et al., 2016). Therefore it is crucial to distinguish episodes of financial contagion from mere interdependences among asset classes/markets, whereby the former is reflected by model parameter instability while evidence of constant parameters should be interpreted as a signal of the latter (see, e.g., Billio and Caporin, 2010). We emphasize that our definition of interdependence and contagion differs from more traditional definitions in the literature (see, among others, Forbes and Rigobon, 2002, and Bacchiochi, 2017) that exploits heteroskedasticity and therefore changes in the contemporaneous cross-correlations of structural shocks, usually in a vector autoregressive model. On the opposite, we study the dynamics of the structural shocks to differentiate mere interdependence from contagion episodes. Intuitively, a constant or smoothly changing VAR coefficient matrix will capture "normal" interdependence among assets through time, whereas a sudden

[1] The way in which stock and (government and corporate) bond markets react to the release of unexpected information or to shocks is likely to differ significantly, so stydying the particular effects in each of these markets, and of course, their comovement and interdependence, is worthy of attention. For instance, Brenner, Pasquariello, and Subrahmantam (2009) document that around macroeconomic announcements, the process of price formation in both stock and bond markets, and their comovement, is driven by fundamentals but with significant asymmetries in their response to unexpected news.

[2] It is now a well established fact that the main source of the recent financial crisis was the turmoil and eventually the ultimate freezing in the subprime mortgage-backed security market (see, e.g., Brunnermeier, 2009; Gorton, 2009). As the property bubble burst after the Summer of 2007, the U.S. financial system immediately suffered from heavy losses that progressively led the US and eventually the global economy into what we now call the "Great Recession". Li, Milne, and Qiu (2016) have presented a model in which contagion and market freezes are caused by uncertainty in financial network structures but there is no doubt that the extent of linkages comes to exceed ordinary interdependence.

and extreme temporal variation of these coefficients during periods of financial stress is considered evidence of contagion.

In our study, we put these definitions at work by first tracking the evolution of three blocks of parameters: those governing the long-term relationships between the endogenous variables in our model, the volatility coefficients of the shocks, and the correlations between such exogenous shocks. Indeed, observing any time-variation in the Bayesian estimates of such blocks of parameters, we proceed to test/detect whether and when any parameter instability occur in the resulting estimates and, thus, obtain a useful measure of contagion. In particular, shifts in (vector) autoregressive coefficients reflect any non-linear effects of variables co-movements that fit Ciccarelli and Rebucci's definition of contagion, while the volatilities of the residuals allow to identify episodes of "apparent" contagion, which are instead simply caused by shifts in exogenous shocks rather than in their transmission mechanisms (see Koop et al., 2009). Moreover, the overall patterns of correlation between pairs of shocks deserves attention in our analysis, because the evolution of these parameters usually becomes particularly erratic during periods of financial distress (for instance, as reported by Casarin, Sartore, and Torzano, 2016). Finally, we also study the instability in the transmission mechanisms of propagation of the shocks by observing how the impulse response functions (henceforth, IRFs) computed under a variety of different models—some featuring contagion and others not—have evolved in the several months following an exogenous shock.

We identify unstable relationships and cross-asset contagion effects using a Bayesian approach, that we prefer over a more traditional frequentist approach, often used in the literature (as in Abbate et al., 2016). There is now a literature that has discussed—often reaching conclusions that favor flexible Bayesian methods—the possibility that frequentist methods may fail to efficiently handle shifting parameters (unless convenient structure is imposed, as in Eickmeier, Lemke, and Marcellino, 2015), a key element to separate out contagion effects from constant-linkage effects in the presence of more variable shocks (see Forbes and Rigobon, 2002). In particular, we propose flexible models that allow for financial instability and cross-asset contagion using Bayesian Vector Autoregressive Models with time varying parameters. We explore several alternative specifications ranging from single-state vector autoregressive (henceforth, VAR) models with constant parameters to fully flexible VAR models where the parameters may vary at each observation. Within the family of Time-Varying-Parameter VAR (TPV-VAR) models, we explore a simple homoskedastic specification, in which the covariance matrix remains constant while the VAR coefficients evolve according to a random walk. Because heteroskedasticity appears to be pervasive across financial data, following Primiceri (2005), we also estimate a TVP-VAR with stochastic volatility (SV) where time variation also affects the covariance matrix. Finally, we extend the range of models entertained to a mixture innovation TVP-VAR model that combines the previous TVP-VAR models with a mixture innovation approach proposed by Gerlach, Carter, and Kohn (2000) and Giordani and Kohn (2008). More precisely, this specification eliminates all restrictions imposed by the standard TVP-VARs and lets "the data speak" about the evolution of all parameters which can either evolve following a random walk or remain constant, depending on sample information, as conveyed by the likelihood function.[3]

All these models are fitted to a multivariate system that includes nine series on the yield series of ABS of different ratings (Aaa for investment- and Bbb for non-investment grade), the repo rate, the 10-year Treasury yield, investment and non-investment, short- and long-term corporate bond yields, and the S&P 500 dividend yield.[4] Of course, basing an empirical exercise of VAR/SV/MI TVP type on as many as nine series poses important computation challenges, although also in this respect our adoption of Monte Carlo Markov Chain (MCMC)-based methods greatly helps with the resulting burden.

Our results provide evidence of intense time variation in both volatilities and correlations of the exogenous shocks. However, according to our posterior estimates, the frequency and the timing of the breaks significantly differ depending on the parameters that are examined. For this reason, the flexible mixture innovation TVP model shows a stronger capability to capture the dynamics of our data compared to standard TVP-VAR models. This evidence is consistent with previous results of Primiceri (2005). Moreover, our results suggest that a few genuine episodes of cross-asset financial contagion especially occur during periods of financial distress. In particular, when we focus on the recent subprime crisis, we find that contagion effects were initially mainly driven by remarkable instability in the endogenous cross-markets relationships, while in the last part of the crisis there is evidence of significant variation also in the scale and co-variation of the exogenous shocks. In other words, 2007 and early 2008 did feature genuine contagion, while—even though in a new regime of stronger interrelationships—the rest of the US crisis (2008 and early 2009) would have instead just featured spillover of large shocks. Finally, our impulse response analysis suggests that, during the period characterized by intense contagion, the propagation of shocks mainly occurs through three of the four channels recently isolated by Longstaff (2010): the flight-to-liquidity, risk premium, and correlated information channels (these are explained in detail in Section 5). We find instead only weak evidence of propagation through a well-known flight-to-quality channel, by which investors would rotate their portfolios off risky assets towards low credit risk ones in times of distress.

Our paper fits at least in three strands of the literature. The literature that falls the closest to our aims is of course a small set of papers specifically concerning contagion, especially with reference to the cross-asset effects observed during the recent financial crisis in the US. For instance, Guo, Chen, and Huang (2011) have studied the collapse of the Collateralized Debt Obligation (CDO) market during the recent financial crisis and show that the default cascade initially affected the mortgage-backed securities and was transmitted to the CDO market and then to the equity market through a contagious negative wealth effect. However, their methods assume that a correlated information

[3]Interestingly, most of these models allow us to capture any deviations in the empirical distribution of returns from normality despite the fact that most models impose the assumption of (multivariate) normality distributed errors at the margin, due to the fact that time-varying means may produce asymmetries and non-zero third-order moments while time-varying variances and covariances always produce fat tails and excess kurtosis (see Marron and Wand, 1992).

[4]As Section 4 explains, in the case of the stock market index we use the (ex-ante) dividend yield instead of standard (ex-post, realized) returns that would also reflect any dividend paid out to align fixed income and repo market quantities (ex-ante rates) with equity ones.

channel would be at work, while in our paper we care for testing contagion vs. interdependence and try and measure the extent of correlated information flows. Also Longstaff (2010) studies the CDO market and analyses the mechanisms through which a negative shock in this market propagates to the other asset classes (Treasuries, corporate bonds, stocks, and the VIX). However, Longstaff potential changes in the linkages across these markets by separately and exogenously analyzing the pre-crisis, the subprime crisis, and the global-crisis period. On the contrary, we are interested in testing the existence of such instability using an encompassing econometric framework in which any shift from simple interdependence to contagion is endogenous. Flavin and Sheenan (2015) have recently extended the analysis in Longstaff (2010) using a Markov Switching Time-Varying Probabilities VAR. Their results surprisingly suggest that cross-assets contagion was weak during the recent subprime crisis, which was mainly driven by stable interdependences across the US markets. Finally, Guo et al. (2011) have studied the contagion effects using a Markov Switching VAR model to analyze the shifting relationships between the stock, credit default, and energy markets. According to their impulse response analysis, contagion effects change significantly depending on the state of the economy, providing much stronger evidence of contagion during the so called “risky/crisis regime”. Of course, a body of empirical work has studied the interdependence and contagion effects during the recent European debt crisis. Bacchiochi (2017) has proposed an heteroskedastic, structural vector autoregression (SVAR) model in which interdependence among European bond markets is captured by the coefficient matrix of the SVAR, while contagion is characterized by changes in the covariance matrix of the shocks during turbulent times. In this study, we identify instead contagion by looking at the sudden changes, regime shifting dynamics of the coefficient matrix of a VAR. Kohonen (2014) has studied the transmission of government default risk in the Eurozone using a SVAR model that allows for structural breaks in the intercepts (which is interpreted as an exogenous change in idiosyncratic country risk), in the autoregressive coefficients (which would capture changes in the interdependence patterns among countries and referred as spillovers effects), and in the contemporaneous cross-correlations, that may capture contagion in times of crisis. Differently from Kohonen (2014), we apply a different identification of contagion episodes and use a more flexible set of TVP models with SV.

A second literature related to our work is the one studying the effects of financial distress on Treasury bond yields, in terms of an alleged flight-to-quality effect. This literature suggests that investors' demand of liquidity changes over time, especially depending on the level of market volatility and supports the belief that, in periods of financial distress, the increase in Treasury bonds' prices is mainly due to the flight-to-liquidity mechanism and counterparty risk.[5] Finally, there is of course a developing literature in which non-linear, multivariate Bayesian methods have been used to account for model instability and contagion in times of crisis, see, among the others, Arakelian and Dellaportas (2012), Bai, Julliard, and Yuan (2012), Casarin, Tronzano, and Sartore (2015), Ciccarelli and Rebucci

[5]Treasury bonds differ from the other financial assets in terms of liquidity and safety (Krishnamurthy and Vissing-Jorgensen , 2012), as they can be easily traded in the market without affecting the price and, at the same time, guarantee a level of credit risk that is lower than any other asset (except for cash).

(2007), and Kaabia, Abid, and Guesmi (2013).

The rest of the paper is organized as follows. In section 2, we describe our multivariate econometric models and the Bayesian estimation methods they require. Section 3 presents the data and descriptive statistics. In section 4, we report our main empirical results. Section 5 evaluates the alternative transmission channels of contagion across asset classes. We conclude in section 6.

2 Methodology

In this section, we describe our general econometric framework and discuss estimation issues and the computation of impulse response functions for the alternative models considered. Similarly to Primiceri (2005) and Koop and Korobilis (2010), we use Time-Varying-Parameter VAR models in order to capture cross-asset contagion. We first discuss the general class of state space models and then present different variants of TVP-VARs along with the MCMC methods required to carry out Bayesian inference. Because our goal remains an applied one, we keep details to a minimum and emphasize instead the different economic insights the alternative models may deliver.

2.1 Normal linear state-space models

A Normal linear state space model is defined by a set of stochastic equations that can be written as

$$\mathbf{y}_t = \mathbf{W}_t\boldsymbol{\delta} + \mathbf{Z}_t\boldsymbol{\beta}_t + \mathbf{u}_t \quad \text{with} \quad \mathbf{u}_t \sim N(\mathbf{0}, \boldsymbol{\Omega}_t) \tag{1}$$

$$\boldsymbol{\beta}_t = \boldsymbol{\Pi}_t\boldsymbol{\beta}_{t-1} + \boldsymbol{\nu}_t \quad \text{with} \quad \boldsymbol{\nu}_t \sim N(\mathbf{0}, \mathbf{Q}_t), \tag{2}$$

where $\mathbf{y}_t$ and $\mathbf{u}_t$ are both $n \times 1$ vectors containing, respectively, the dependent variables and the error terms in the measurement equation (1) that captures how (say) asset yields depend on vector of predetermined variables, some of them interacted with fixed coefficients $\boldsymbol{\delta}$, and others with time-varying coefficients, $\boldsymbol{\beta}_t$. $\boldsymbol{\delta}$ is a $p_0 \times 1$ vector of constant coefficients that map the known $n \times p_0$ matrix of explanatory variables $\mathbf{W}_t$ into the dependent variables; while $\boldsymbol{\beta}_t$ is a $m \times 1$ vector of time varying coefficients that multiplies the known $n \times m$ matrix of explanatory variables $\mathbf{Z}_t$. We assume that also the $m \times m$ matrix $\boldsymbol{\Pi}_t$ featured by the state equation (2) is predetermined at time t and that the vector of error terms, $\mathbf{u}_t$ and $\boldsymbol{\nu}_t$, are independent over time and from one another. However, $\boldsymbol{\Omega}_t$ and $\mathbf{Q}_t$ need not be diagonal (or scalar), which means that at time t, the shocks to both measurement and state equations may be cross-correlated. Of course, when $\boldsymbol{\nu}_t = \mathbf{0}$, and $\boldsymbol{\Pi}_t = \mathbf{I}_m$ $\forall t$, (3)-(4) becomes a single-state, constant coefficient VARX(q) that may be still estimated using Bayesian methods, e.g., as in Mumtaz and Surico's (2009) analysis of international shock transmission in large-scale VARs. In spite of the multivariate normality assumption in (1)-(2) for the shocks, time-varying parameter models are flexible enough to accommodate highly non-normal shapes of the distribution of the variables of interest: time-varying conditional means may produce asymmetries and non zero third-order moments and time-varying variances/covariances always produce fat tails and excess kurtosis (see Marron and

Wand, 1992).

Broadly speaking, in a Bayesian framework, our goal is to obtain from a given sample $\mathbf{y}^T \equiv [\mathbf{y}_1' \ \mathbf{y}_2' \ ... \ \mathbf{y}_T']'$ the posterior densities of the unknown coefficients, $\boldsymbol{\delta}$, $\boldsymbol{\beta}_{t+1}$, $\boldsymbol{\Pi}_t$, $\mathbf{Q}_t$, and $\boldsymbol{\Omega}_t$, to be able to draw inferences on them and estimate meaningful quantities of interest, such as impulse response functions. In a Bayesian set up, this is performed by first setting up adequate (economically sensible priors) on the matrices $\boldsymbol{\delta}$, $\boldsymbol{\beta}_{t+1}$, $\boldsymbol{\Pi}_t$, $\mathbf{Q}_t$, and $\boldsymbol{\Omega}_t$, and then combining these with the evidence contained in the data, as summarized by the likelihood function. This is obtained by applying MCMC algorithms explained in the following. However, before proceeding in this way and also to gain some concreteness, we present and discuss a few specializations of the general framework (1)-(2).

2.2 Homoskedastic TVP-VAR

The homoskedastic version of the TVP-VAR model is a restricted version of the Normal linear state-space model (1)-(2), in which the covariance matrices are both assumed to be constant over time:

$$\mathbf{y}_t = \mathbf{Z}_t\boldsymbol{\beta}_t + \mathbf{u}_t \quad \text{with} \quad \mathbf{u}_t \sim N(\mathbf{0}, \boldsymbol{\Omega}) \tag{3}$$

$$\boldsymbol{\beta}_{t+1} = \boldsymbol{\beta}_t + \boldsymbol{\nu}_{t+1} \quad \text{with} \quad \boldsymbol{\nu}_{t+1} \sim N(\mathbf{0}, \mathbf{Q}), \tag{4}$$

where, u_t and ν_{t+s} are assumed to be independent from one another for any s and t. Importantly, the model is also specialized to a state equation that follows a multi-dimensional random walk, $E[\boldsymbol{\beta}_{t+1}] = \boldsymbol{\beta}_t$; moreover, any exogenous variable has been excluded and this turns (1)-(2) into a VAR(q) model if one sets $\mathbf{Z}_t \equiv [\mathbf{y}_{t-1}' \ \mathbf{y}_{t-2}' \ ... \ \mathbf{y}_{t-q}']$ and $m = n^2q$.[6]

Economically, a TVP-VAR model is one in which, because of the homoskedasticity, the scale of the shocks is constant over time (and controlled by the fixed matrix $\boldsymbol{\Omega}$), while each period is characterized by structural instability of the beta coefficients mapping past values of the yields into future ones. This is caused by the fact that even though $E[\boldsymbol{\beta}_{t+1}] = \boldsymbol{\beta}_t$, $\Pr(\boldsymbol{\beta}_{t+1} = \boldsymbol{\beta}_t) = 0$ except for a null measure event (i.e., it is just very hard for this to happen). This implies that cross-market instabilities will be pervasive, even though their scale may be modest, when the scale of the associated shocks ($\mathbf{Q}$) turns out to be "small". Therefore (3)-(4) is a model of extremely frequent instabilities, entirely driven by small but continuous shifts in the values of the coefficients that map past yields into forecasts of the future. Even though, this may seem extreme and unrealistic, we shall let the data judge about the empirical merits of this model.

The state equation (4) naturally provides a time-varying prior for $\boldsymbol{\beta}_{t+1}$, as it implies $\boldsymbol{\beta}_{t+1}|\boldsymbol{\beta}_t, \mathbf{Q} \sim N(\boldsymbol{\beta}_t, \mathbf{Q})$, from which the following joint prior for the overall sequence $\boldsymbol{\beta}^T \equiv [\boldsymbol{\beta}_0' \ \boldsymbol{\beta}_1' \ \boldsymbol{\beta}_2' \ ... \ \boldsymbol{\beta}_T']'$ can be derived:

$$p(\boldsymbol{\beta}^T|\mathbf{Q}) = p(\boldsymbol{\beta}_0)\prod_{t=1}^{T} f(\boldsymbol{\beta}_t|\boldsymbol{\beta}_{t-1}, \mathbf{Q}), \tag{5}$$

[6] In this case, one obtains: $\mathbf{y}_t = \mathbf{c}_t + \mathbf{B}_{1,t}\mathbf{y}_{t-1} + \mathbf{B}_{2,t}\mathbf{y}_{t-2} + ... + \mathbf{B}_{q,t}\mathbf{y}_{t-q} + \mathbf{u}_t$ with $\mathbf{u}_t \sim N(\mathbf{0}, \boldsymbol{\Omega})$.Stationarity of the VAR is enforced through a simple rejection method, i.e., combinations of coefficients sampled during the Gibbs procedure that imply a violation of stationarity are dealt with by rejecting the entire corresponding, to be re-sampled.

where $f(\cdot)$ represents the normal density function. Additionally, we also select the following priors:

$$\boldsymbol{\beta}_0 \sim N(\widehat{\boldsymbol{\beta}}_{OLS}, 4Cov(\widehat{\boldsymbol{\beta}}_{OLS})) \qquad \boldsymbol{\Omega}^{-1} \sim W(\underline{\mathbf{S}}^{-1}, \underline{\upsilon}) \qquad \mathbf{Q}^{-1} \sim W(\underline{\mathbf{Q}}^{-1}, \underline{\upsilon}_Q), \tag{6}$$

where $W(\mathbf{X}, \psi)$ indicates a matrix Wishart distribution for the random matrix $\mathbf{X}$, we set $\underline{\mathbf{S}} = \mathbf{I}_n$, $\underline{\upsilon} = n+1$, $\underline{\mathbf{Q}} = 0.0001\tau Cov(\widehat{\boldsymbol{\beta}}_{OLS})$, and $\underline{\upsilon}_Q = \tau$. The coefficient τ is the number of pre-sample observations used to calibrate these (empirical) priors, for instance by estimating by standard OLS $\widehat{\boldsymbol{\beta}}_{OLS}$ and $Cov(\widehat{\boldsymbol{\beta}}_{OLS})$. In our applications, we set $\tau = 40$ as typical of many papers in the empirical finance and macroeconomics literatures (see, e.g., Koop and Korobilis, 2010). Under the assumption of independent priors, $p(\boldsymbol{\beta}^T, \boldsymbol{\Omega}, \mathbf{Q}) = p(\boldsymbol{\beta}^T)\ p(\boldsymbol{\Omega})\ p(\mathbf{Q})$, we apply the Gibbs sampler to sequentially draw from the conditional posteriors $p(\mathbf{Q}^{-1}|\mathbf{y}^T, \boldsymbol{\beta}^T)$, $p(\boldsymbol{\Omega}^{-1}|\mathbf{y}^T, \boldsymbol{\beta}^T)$, and $p(\boldsymbol{\beta}^T|\mathbf{y}^T, \boldsymbol{\Omega}, \mathbf{Q})$ following the iterative algorithm:

$$\begin{aligned}
\boldsymbol{\Omega}^{-1}|\mathbf{y}^T, \boldsymbol{\beta}^T &\sim W(\overline{\mathbf{S}}^{-1}, \overline{\upsilon}), \qquad \overline{\upsilon} = \underline{\upsilon} + T \qquad \overline{\mathbf{S}} = \underline{\mathbf{S}} + \sum_{t=1}^{T}(\mathbf{y}_t - \mathbf{Z}_t\boldsymbol{\beta}_t)(\mathbf{y}_t - \mathbf{Z}_t\boldsymbol{\beta}_t)' \\
\mathbf{Q}^{-1}|\mathbf{y}^T, \boldsymbol{\beta}^T &\sim W(\overline{\mathbf{Q}}^{-1}, \overline{\upsilon}_Q), \qquad \overline{\upsilon}_Q = \underline{\upsilon}^Q + T \qquad \overline{\mathbf{Q}} = \underline{\mathbf{Q}} + \sum_{t=1}^{T}(\boldsymbol{\beta}_{t+1} - \boldsymbol{\beta}_t)(\boldsymbol{\beta}_{t+1} - \boldsymbol{\beta}_t)'
\end{aligned} \tag{7}$$

At this point, for given posteriors for $\boldsymbol{\Omega}$ and $\mathbf{Q}$, we perform the posterior simulation of $\boldsymbol{\beta}^T$ applying the methods proposed by Carter and Kohn(1994).

2.3 TVP VAR with Stochastic Volatility

In financial applications, it is preferable to relax the assumption of homoskedasticity and allow $\mathbf{u}_t$ to be conditionally IID normal $(\mathbf{0}, \boldsymbol{\Omega}_t)$, in which the covariance matrix of error terms $\boldsymbol{\Omega}_t$ is time-varying. In order to address this issue, we expand the set of VAR models used in our estimations to include stochastic variations of TVP-VAR.[7] We start from a simple univariate stochastic volatility, in which residuals are modeled as follows:

$$y_{i,t} = \mathbf{W}_{i,t}\boldsymbol{\delta}_i + \exp\left(\frac{h_{i,t}}{2}\right)u_{i,t} \quad \text{with} \quad u_{i,t} \sim N(0,1) \qquad i = 1, 2, ..., n \tag{8}$$

$$h_{i,t+1} = \mu_i + \phi_i(h_{i,t} - \mu_i) + \eta_{i,t} \quad \text{with} \quad \eta_{i,t} \sim N(0, \sigma^2_{i,\eta}) \tag{9}$$

where $u_{i,t}$ and $\eta_{i,t+s}$ independent from one another for all s and t.[8] Equations (8) and (9) can be interpreted as a state space model where, in contrast to the linear model defined by (1)-(2), the dependent variable $y_{i,t}$ is not a simple linear function of the state, here defined by the scale variable $h_{i,t}$. Following Koop and Korobilis (2010), it is interesting to note that $h_{i,t}$ corresponds to the log-standard deviation of $y_{i,t}$. However, to allow the error term to be normally distributed, we will refer

[7] We refer to West and Harrison (1997) and Kim and Nelson (1999) for a complete description of the wide class of nonlinear state space models with SV. The non-linearity derives from the fact that it is non-linear transformations of the data that pin down the process followed by the covariance matrix of the shocks.

[8] Of course, although technically interesting, a model in which $y_{i,t} = \mathbf{W}_{i,t}\boldsymbol{\delta}_i + \mathbf{Z}_{i,t}\boldsymbol{\beta}_{i,t} +$ exp $\left(\frac{h_{i,t}}{2}\right)u_{i,t}$ with $u_{i,t} \sim N(0,1)$, represents a straightforward extension of the model in (8).

to equation (9) as the log-volatility process.

Of course, a contagion analysis would enormously benefit from the adoption of models in which not only the volatilities, but also the covariances of yields could be stochastic, to capture the fact that not only shocks but also their (linear) association may dynamically evolve over time. As discussed in Section 1, to draw a clear distinction between volatility spillovers and contagion is a key task, as emphasized since the seminal paper by Forbes and Rigobon (2002). Following Primiceri (2005), we extend the model in (8)-(9) to a multivariate framework; this can be done rather easily if one accepts that—because in high-frequency financial data volatility persistence tends to be high—setting $\phi_i = 1$ as in Primiceri's work, to imply a random walk specification for the state equation. In particular, the TVP-VAR model with stochastic volatility is given by:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{c}_t + \mathbf{B}_{1,t}\mathbf{y}_{t-1} + \mathbf{B}_{2,t}\mathbf{y}_{t-2} + ... + \mathbf{B}_{q,t}\mathbf{y}_{t-q} + u_t \quad \text{with} \quad u_t \sim N(\mathbf{0}, \mathbf{\Omega}_t) &(10)\\ \mathbf{A}_t\mathbf{\Omega}_t\mathbf{A}_t' &= \mathbf{\Sigma}_t\mathbf{\Sigma}_t' &(11)\end{aligned}$$

where $\mathbf{A}_t\mathbf{\Omega}_t\mathbf{A}_t'$ represents the lower triangular decomposition of the covariance matrix $\mathbf{\Omega}_t$ introduced by Primiceri to increase the efficiency of the estimates, where $\mathbf{\Sigma}_t$ is a diagonal matrix with standard deviations of the n assets on its main diagonal.

Thus, we can re-write the model in (10)-(11) in compact form as:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{Z}_t'\boldsymbol{\beta}_t + \mathbf{A}_t^{-1}\mathbf{\Sigma}_t\boldsymbol{\varepsilon}_t \qquad \text{with } \boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \mathbf{I}_n) &(12)\\ \mathbf{Z}_t' &\equiv \mathbf{I}_n \otimes [1 \ \mathbf{y}_{t-1}... \ \mathbf{y}_{t-1}], \ \boldsymbol{\beta}_t \equiv [\mathbf{c}_t \ \mathbf{B}_{1,t} \ ...\mathbf{B}_{q,t}]. &(13)\end{aligned}$$

Following Primiceri (2005), we stack the unrestricted elements of the matrices $\mathbf{A}_t$ and $\mathbf{\Sigma}_t$ in the vectors $\mathbf{a}_t$ and $\boldsymbol{\sigma}_t$, respectively, and assume that the corresponding coefficients follow the process:

$$\begin{aligned} \boldsymbol{\beta}_t &= \boldsymbol{\beta}_{t-1} + \boldsymbol{\nu}_t\\ \mathbf{a}_t &= \mathbf{a}_{t-1} + \boldsymbol{\zeta}_t\\ \ln(\boldsymbol{\sigma}_t) &= \ln(\boldsymbol{\sigma}_{t-1}) + \boldsymbol{\eta}_t, &(14)\end{aligned}$$

where $\ln(\boldsymbol{\sigma}_t)$ indicates the element-wise natural log of the elements of $\boldsymbol{\sigma}_t$. Moreover, without loss of generality, we set the innovation terms $(\boldsymbol{\varepsilon}_t, \boldsymbol{\nu}_t, \boldsymbol{\zeta}_t, \boldsymbol{\eta}_t)'$ to be jointly normally distributed with

$$\mathbf{V} = Cov\begin{pmatrix} \boldsymbol{\varepsilon}_t \\ \boldsymbol{\nu}_t \\ \boldsymbol{\zeta}_t \\ \boldsymbol{\eta}_t \end{pmatrix} = \begin{bmatrix} \mathbf{I}_n & \mathbf{O} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{Q} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{S} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \mathbf{W} \end{bmatrix},$$

and **Q**, **S,** and **W** positive definite matrices.[9] In an economic perspective, also the model in (10)-(11)

[9] **S** is assumed to have a block diagonal structure in order to further increase the efficiency of our estimation procedure. All the zero elements could be replaceed by non-zero blocks; however, we opt for this specification in order to reduce the

forces all parameters, in this case also the variances and (indirectly, through the $\mathbf{\Omega}_t = \mathbf{A}_t'\mathbf{\Sigma}_t\mathbf{\Sigma}_t'\mathbf{A}_t$ transformation) covariances, to change in correspondence to all point in the sample. Of course, by making the covariance matrix of all shocks $\mathbf{V}$ sufficiently small, it is then possible to capture situations in which parameters generally evolve smoothly with small period changes, apart from occasional big but infrequent shocks hitting them.

As in the homoskedastic TVP-VAR case, we obtain OLS estimates of $\boldsymbol{\beta}_0$ from a training sample of $\tau = 40$ initial observations and use them to set the prior distributions as

$$\begin{aligned} \boldsymbol{\beta}_0 &\sim N(\widehat{\boldsymbol{\beta}}_{OLS}, 4Cov(\widehat{\boldsymbol{\beta}}_{OLS})), \qquad \mathbf{A}_0 \sim N(\hat{\mathbf{A}}_{OLS}, 4Cov(\hat{\mathbf{A}}_{OLS})), \\ \ln \boldsymbol{\sigma}_0 &\sim N(\ln \hat{\sigma}_{OLS}, 4\mathbf{I}_n), \qquad \mathbf{Q} \sim IW(k_Q^2 \tau Cov(\widehat{\boldsymbol{\beta}}_{OLS}), \tau), \\ \mathbf{W} &\sim IW(k_W^2 (n+1)\mathbf{I}_n, (n+1)), \qquad \mathbf{S}_{ii} \sim IW(k_S^2 (i+1) Cov(\hat{\mathbf{A}}_{ii,OLS}), (i+1)), \end{aligned} \tag{15}$$

where $IW(\mathbf{X}, \psi)$ indicates a matrix inverse Wishart distribution for the random matrix $\mathbf{X}$, $k_Q = 0.01$, $k_S = 0.1$, and $k_w = 0.01$; $\mathbf{S}_{ii}$ and $\hat{\mathbf{A}}_{ii,OLS}$ denote the ith block of $\mathbf{S}$ and the correspondent block of $\hat{\mathbf{A}}_{OLS}$, respectively. $\hat{\mathbf{A}}_{OLS}$ can be computed from the relationship $\mathbf{A}Cov(\widehat{\mathbf{u}}_t^{OLS})\mathbf{A}' = diag(\hat{\boldsymbol{\sigma}}_{OLS}^2)$. Based on the selected priors and conditionally on the data observations, the Gibbs sampler performs posterior simulations sequentially drawing the VAR coefficients $\boldsymbol{\beta}^T$, the simultaneous relations $\mathbf{A}^T \equiv [\mathbf{A}_0 \; \mathbf{A}_1 \; \mathbf{A}_2 \; ... \; \mathbf{A}_T]'$, the volatilities $\mathbf{\Sigma}^T \equiv [\mathbf{\Sigma}_0 \; \mathbf{\Sigma}_1 \; \mathbf{\Sigma}_2 \; ... \; \mathbf{\Sigma}_T]'$, and the hyperparameters contained in $\mathbf{V}$. As for the previous specification, the algorithm of Carter and Kohn (1994) is applied to draw $\boldsymbol{\beta}^T$ from the posterior $p(\boldsymbol{\beta}^T | \mathbf{y}^T, \mathbf{A}^T, \mathbf{\Sigma}^T, \mathbf{V})$.

2.4 TVP VAR with Mixture Innovations

The pioneering model presented by Primiceri (2005) has marked a crucial turning point in the TVP VAR literature, as it first introduced time variation both in the transmission mechanisms of past shocks and in the covariance matrix of the shocks. However, allowing the parameters to change at each new observation, the two types of TVP models presented so far are built on the restrictive assumption that $T-1$ small but "persistently repeated" breaks will occur in a sample of size T. In order to address this limitation, we add to the TVP specification presented in the previous sections also the mixture innovation approach of Gerlach et al. (2000) and Giordani and Kohn (2008), which allows to determine the number, as well as the timing, of any changes in model parameters directly from the data. This a model of infrequent, possibly rare, parameter shifts that can indeed trigger periods of altered co-movements that we normally dub as contagion episodes.

Following Koop et al. (2009), we extend Primiceri (2005) by introducing within the framework (10)-(11), for any $t = 1, 2, \ldots, T$, three Markov random variables $\mathbf{K}_t \equiv [K_{1t} \; K_{2t} \; K_{3t}]'$ that respectively controls for breaks in the VAR coefficients ($\boldsymbol{\beta}_t$), in the simultaneous relations that drive the mapping from variances to covariances between pairs of shocks ($\mathbf{A}_t$), and in the variances of the shocks ($\mathbf{\Sigma}_t$).

number of parameters in the model and allow for a structural interpretation of the innovations.

This occurs according to the dynamics

$$\begin{aligned}
\boldsymbol{\beta}_t &= \boldsymbol{\beta}_{t-1} + K_{1t}\boldsymbol{\nu}_t \\
\mathbf{a}_t &= \mathbf{a}_{t-1} + K_{2t}\boldsymbol{\zeta}_t \\
\ln(\boldsymbol{\sigma}_t) &= \ln(\boldsymbol{\sigma}_{t-1}) + K_{3t}\boldsymbol{\eta}_t,
\end{aligned} \tag{16}$$

where $K_{it} \in \{0, 1\}$ for $t = 1, 2, 3$. Crucially the breaks are not restricted from occurring at the same time. Thus, at any time t, the parameter vectors $\boldsymbol{\beta}_t, \mathbf{a}_t$ and $\ln(\boldsymbol{\sigma}_t)$ can either evolve according to a random walk (if one or more among K_{1t}, K_{2t}, and K_{3t} equal 1) or stay constant (if one or more among K_{1t}, K_{2t}, and $K_{3t} = 0$). If the posterior density of $\mathbf{K}_t$ puts considerable mass on $[0\ 0\ 0]'$, then most of the time it will be $\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1}$, $\mathbf{a}_t = \mathbf{a}_{t-1}$, and $\ln(\boldsymbol{\sigma}_t) = \ln(\boldsymbol{\sigma}_{t-1})$, and both the scale, the covariance, and the transmission structure of shocks will be constant between two consecutive periods, like in a plain-vanilla VAR(q) model. However, as soon as one of the Markov "switching" variables is triggered into taking a unit value, a break in parameters will occur, for instance: $\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\nu}_t$, $\mathbf{a}_t = \mathbf{a}_{t-1} + \boldsymbol{\zeta}_t$, but $\ln(\boldsymbol{\sigma}_t) = \ln(\boldsymbol{\sigma}_{t-1})$, which is the case in which the mechanism of transmission of past shocks undergoes a shift that depends on the realization of $\boldsymbol{\nu}_t$, the covariance structure of the shocks undergoes a shift the depends on the realization of $\boldsymbol{\zeta}_t$, but the scale of all shocks to yields remains constant. In fact five interesting cases are possible:

1. $\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1}$, $\mathbf{a}_t = \mathbf{a}_{t-1}$, and $\ln(\boldsymbol{\sigma}_t) = \ln(\boldsymbol{\sigma}_{t-1})$: no contagion, no volatility spillover, just normal interdependence.
2. $\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\nu}_t$, $\mathbf{a}_t = \mathbf{a}_{t-1}$, and $\ln(\boldsymbol{\sigma}_t) = \ln(\boldsymbol{\sigma}_{t-1})$: contagion purely driven by altered transmission mechanism.
3. $\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1}$, $\mathbf{a}_t = \mathbf{a}_{t-1} + \boldsymbol{\zeta}_t$, and $\ln(\boldsymbol{\sigma}_t) = \ln(\boldsymbol{\sigma}_{t-1})$: no contagion, just altered interdependence of shocks due to volatility spillover.
4. $\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1}$, $\mathbf{a}_t = \mathbf{a}_{t-1}$, but $\ln(\boldsymbol{\sigma}_t) = \ln(\boldsymbol{\sigma}_{t-1}) + \boldsymbol{\eta}_t$: no contagion and normal interdependence, just altered volatility that however spreads in normal ways, typical of periods of financial distress.
5. $\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\nu}_t$, $\mathbf{a}_t = \mathbf{a}_{t-1} + \boldsymbol{\zeta}_t$, but $\ln(\boldsymbol{\sigma}_t) = \ln(\boldsymbol{\sigma}_{t-1})$: both contagion and volatility spillover at work, from altered interdependence.[10]

As noted in Koop et al. (2009), this is a very flexible model, since it nests both the Homoskedastic TVP-VAR in (3)-(4) (this occurs when $K_{1t} = 1, K_{2t} = K_{3t} = 0$ for all t) and the TVP VAR with stochastic volatility ($K_{1t} = K_{2t} = K_{3t} = 1$ for any t), which can be used to fit small and gradual changes in the parameters. But, at the same time, it offers the possibility to use a more parsimonious model, able to deal with an unknown number of breaks, by simply relaxing any restriction on the

[10] The case of $\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1}$, $\mathbf{a}_t = \mathbf{a}_{t-1} + \boldsymbol{\zeta}_t$, and $\ln(\boldsymbol{\sigma}_t) = \ln(\boldsymbol{\sigma}_{t-1}) + \boldsymbol{\eta}_t$; $\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\nu}_t$, $\mathbf{a}_t = \mathbf{a}_{t-1}$, and $\ln(\boldsymbol{\sigma}_t) = \ln(\boldsymbol{\sigma}_{t-1}) + \boldsymbol{\eta}_t$ are of course possible. They occur when a crisis causes both altered interdipendence and/or contagion, as well as higher volatility than in normal times.

evolution of the states. In Section 4, we have estimated all these variants of the baseline model also to try and economize over the complexity implied.

One important restriction of the general mixture model in Koop et al. (2009) and Primiceri (2005) is represented by what we call the Cogley-Sargent variant (2005), i.e., when in $\mathbf{K}_t$, $K_{2t} = 0\ \forall t$:

$$\begin{aligned}\boldsymbol{\beta}_t &= \boldsymbol{\beta}_{t-1} + K_{1t}\boldsymbol{\nu}_t \\ \mathbf{a}_t &= \mathbf{a}_{t-1} = \bar{\mathbf{a}} \\ \ln(\boldsymbol{\sigma}_t) &= \ln(\boldsymbol{\sigma}_{t-1}) + K_{3t}\boldsymbol{\eta}_t. \end{aligned} \tag{17}$$

In this model, the scale of shocks, is subject to possibly large but infrequent shifts driven by K_{3t}; moreover, the way shocks propagate is subject to breaks, which we have called contagion episodes, governed by K_{1t}. However, the simultaneous interdependence across shocks, as measured by $\mathbf{A}_t$ are non-zero (in fact, they may even be massive and exceed in importance those explained by the VAR matrices collected in $\boldsymbol{\beta}_t$), but are restricted to be constant over time. Obviously, this is just alternative 4 above, elevated to the rank of separate model because the restriction is imposed ex-ante, as if it were a infinitely precise prior.

Model estimation is performed including few additional specifications in the Gibbs Sampler that are now getting standard in the literature, see for instance Gerlach et al. (2000).[11] In particular, we need to define the prior distribution of the vector $\mathbf{K}$ along with a methodology to get its posterior density by simulation. Following Koop et al. (2009), we use a Markov hierarchical prior $p(K_{it} = 1) = p_i$ for $i = 1, 2, 3$, where the (conditionally) conjugate Beta prior $Beta(\underline{\beta}_{1i}, \underline{\beta}_{2i})$ is used for the change-point probabilities p_1, p_2, and p_3. As a result, the conditional posterior for p_i , that will be used throughout the Gibbs Sampler iterations, is $Beta(\overline{\beta}_{1i}, \overline{\beta}_{2i})$ with $\overline{\beta}_{1i} = \underline{\beta}_{1i} + \sum_{t=1}^{T} K_{it}$, $\overline{\beta}_{2i} = \underline{\beta}_{2i} + T - \sum_{t=1}^{T} K_{it}$. Based on this prior, that assumes breakpoints in $\boldsymbol{\beta}_t$, $\mathbf{a}_t$, and $\ln(\boldsymbol{\sigma}_t)$ to be independent of one another, we can individually generate K_{1t}, K_{2t}, K_{3t} using the efficient algorithm of Gerlach et al. (2000). Note that the fact that breaks in VAR coefficients, covariances of the shocks, and variances are assumed to be independent in the priors does not imply that they will be in the posteriors and hence in the inferred estimates. If the data were to contain evidence of *co-breaking patterns* across different coefficients, the likelihood function will carry this information and posterior estimates of the historical path followed by $\mathbf{K}^T \equiv [\mathbf{K}_1^T\ \mathbf{K}_2^T\ \mathbf{K}_3^T]$ will reflect this pattern.

In practice, the priors are set whenever possible to be identical to those used for the standard TVP-VAR model of Sections 2.2-2.3. As for the additional mixture innovation parameters, we opt for an informative "few-breaks" Beta prior $B(\underline{\beta}_{1j}, \underline{\beta}_{2j})$ with $\underline{\beta}_{1j} = 0.1$ and $\underline{\beta}_{2j} = 50$, for all $j = 1, 2, 3$. According to the properties of the Beta distribution, this prior implies a break probability of 0.02 only. This guarantees that any evidence of frequent breaks comes not from our willingness to detect

[11] To use the Gibbs sampler within the new TVP-VAR specification, we only need to make a few minor adjustments. Firstly, all the sequential draws in the MCMC algorithm must be taken conditionally on $\mathbf{K}$. Moreover, the conditional posterior for the degrees of freedom of $\mathbf{Q}$, $\mathbf{S}$ and $\mathbf{W}$ shall not depend on the total sample size T but on the sum of the corresponding (that are associated to breaks in the underlying shocks) values of $\mathbf{K}$ (i.e., $\sum_{t=1}^{T} K_{it}$ for $i = 1, 2, 3$)

them, but from the data, through the likelihood function.

2.5 Markov Switching VAR

Finally, because this model has often appeared in the literature on cross-asset contagion (see, e.g., Guo et al., 2011; Guidolin and Pedio, 2017; Flavin and Sheenan, 2015; Connolly, Stivers, and Sun, 2007), we adopt as an additional benchmark also a four-state Markov switching VAR (MSVAR) model, that can be written as (12)-(13) supplemented by the further specification that:

$$\begin{aligned} \boldsymbol{\beta}_t &= (1 - S_{1t})\boldsymbol{\beta}_0 + S_{1t}\boldsymbol{\beta}_1 \\ \mathbf{a}_t &= (1 - S_{2t})\mathbf{a}_0 + S_{2t}\mathbf{a}_1 \\ \ln(\boldsymbol{\sigma}_t) &= (1 - S_{2t})\ln(\boldsymbol{\sigma}_0) + S_{2t}\ln(\boldsymbol{\sigma}_0). \end{aligned} \tag{18}$$

$S_{1t} = 0, 1$ is a first-order, two-state, homogeneous, irreducible, and ergodic Markov chain variable that drives regime shifts in the VAR coefficient matrices; $S_{2t} = 0, 1$ is a first-order, two-state, homogeneous, irreducible, and ergodic Markov chain variable that drives regime shifts in variance and covariances of the shocks.[12] Note that in line with most of the literature, we assume switching in variances and covariances to occur simultaneously. Because we assume that S_{1t} and S_{2t} are independent, the resulting MSVAR model is effectively a four-state one, which seems adequate both to describe the financial data from as many as nine different markets, and to capture the fact that regime shifts in the conditional mean and the conditional variance and covariance function may occur at different times. In fact, the very notion of contagion on top of pure interdependence, is that contagion would be triggered by a change in the way in which shocks are transmitted across markets, given any shift in variances and covariances of the very shocks. If $S_{1t} = 1$ refers to contagion periods (labelling regimes in MSVAR models is arbitrary), then the vector of VAR coefficients $\boldsymbol{\beta}_1$ will imply stronger linkages than the "normal regime" coefficients, $\boldsymbol{\beta}_0$. If instead turbulent periods are identified by $S_{2t} = 1$, then it is possible for them to be accompanied by abnormally high correlation between shocks, presumably identified by a vector $\mathbf{a}_1$ that differs from $\mathbf{a}_0$. Finally, note that while in a mixture innovation TVP-VAR it is possible for multiple regime shifts (breaks) to occur over time without having a recurring nature because $\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\nu}_t$, $\mathbf{a}_t = \mathbf{a}_{t-1} + \boldsymbol{\zeta}_t$, or $\ln(\boldsymbol{\sigma}_t) = \ln(\boldsymbol{\sigma}_{t-1}) + \boldsymbol{\eta}_t$ simply take the coefficient to new values that purely depend on the realization of the shock vectors $\boldsymbol{\nu}_t$, $\boldsymbol{\zeta}_t$, and $\boldsymbol{\eta}_t$, the recurring regimes in a MSVAR operate in a simpler way, by taking $\boldsymbol{\beta}_t$ from $\boldsymbol{\beta}_0$ to $\boldsymbol{\beta}_1$ and/or $\mathbf{a}_t$ and $\ln(\boldsymbol{\sigma}_t)$ from $\mathbf{a}_0$ and $\ln(\boldsymbol{\sigma}_0)$ to $\mathbf{a}_0$ and $\ln(\boldsymbol{\sigma}_1)$. Although this may appear over-simplistic compared to mixture TVP-VARs, MSVARs capture in a very simple way the idea that coefficients will be unstable but with rather infrequent shifts but make it possible a much simpler estimation approach vs. models characterized by shifts governed by $\mathbf{K}_t \equiv [K_{1t}\ K_{2t}\ K_{3t}]'$.

[12] The first-order nature and homogeneity property guarantee that each Markov chain is characterized by constant transition probabilities. Effectively, these are four parameters: $p^1_{00} \equiv \Pr(S_{1t} = 0|S_{1t-1} = 0)$, $p^1_{11} \equiv \Pr(S_{1t} = 1|S_{1t-1} = 1)$, $p^2_{00} \equiv \Pr(S_{2t} = 0|S_{2t-1} = 0)$, and $p^2_{11} \equiv \Pr(S_{2t} = 1|S_{21t-1} = 1)$. These populate a 8×8 transition matrix obtained as the Kronecker's product between the 2×2 transition matrix for S_1 and the 2×2 transition matrix for S_2.

Estimation is performed by setting the initial values of the parameters to equal the MLE estimates gotten using the standard expectation-maximization algorithm and then sequentially going through the steps of the Gibbs sampler. The initial state is also set to match the vector of ergodic state probabilities of the process. Kim and Nelson (1999) and Fruhwirth-Schnatter (2006) provide a step-by-step explanation of the algorithm and heavy doses of important details necessary to its implementation.

2.6 Generalized Impulse Response Functions

We conclude this methodological section with a brief survey of the estimation process applies to the impulse response functions (henceforth, IRFs) for the various TVP-VAR models introduced above. To estimate IRFs, as a first step one needs to move from (12)-(13) to the structural form of the model:

$$\mathbf{y}_t = \mathbf{Z}_t' \boldsymbol{\beta}_t + \boldsymbol{\Upsilon}_t \boldsymbol{\varepsilon}_t \qquad \text{with } \boldsymbol{\Upsilon}_t = \mathbf{A}_t^{-1} \boldsymbol{\Sigma}_t. \tag{19}$$

The restrictions implicit in $\boldsymbol{\Upsilon}_t = \mathbf{A}_t^{-1}\boldsymbol{\Sigma}_t$ provide the identifying assumptions to isolate the structural shocks and the elements of $\boldsymbol{\varepsilon}_t$ correspond to the structural form errors. Thus, at each MCMC iteration, we transform the obtained draws for $\mathbf{a}_t$ (hence, $\mathbf{A}_t$) and $\ln(\boldsymbol{\sigma}_t)$ (hence, $\boldsymbol{\Sigma}$) to obtain a posterior draw for $\boldsymbol{\Upsilon}_t$ and, hence, impulse responses calculated in standard ways, when the structural shocks obey the identification provided by $\boldsymbol{\Upsilon}_t$. Following Primiceri (2005), we assume that $\boldsymbol{\Upsilon}_t$ imposes at least $n(n-1)/2$ restrictions in order to guarantee identification. More precisely, the identification process requires to first estimate the reduced form model and then derive the matrix $\boldsymbol{\Upsilon}_t$ such that $\boldsymbol{\Upsilon}_t\boldsymbol{\Upsilon}_t' = \boldsymbol{\Omega}_t$, for $t = 1, \ldots, T$ and for every draw of $\boldsymbol{\Omega}_t$.[13]

Because both the posterior inferences for the coefficients in $\boldsymbol{\beta}_t$ and for $\boldsymbol{\Upsilon}_t$ may change over time, in TVP-VAR models the resulting IRFs have shapes that in general are expected to change over time. Generally speaking, for a nonlinear model, the H-horizon/step impulse response function to a shock in variable s ($s = 1, 2, ..., n$) is defined as

$$IRF_{t,H}(\mathbf{S}^t) = E[y_{i,t+H}|\mathbf{y}^t, \mathbf{S}^t, \varepsilon_t + \boldsymbol{\nabla}_{\boldsymbol{\varepsilon}}] - E[y_{i,t+H}|\mathbf{y}^t, \mathbf{S}^t, \varepsilon_t] \qquad i = 1, 2, ..., n, \tag{20}$$

where $\mathbf{y}^t$ corresponds to all the data available up to time t, $\mathbf{S}^t$ explicitly appears to accommodate the MSVAR case in which shifts in parameters display memory because they are persistent, and $\boldsymbol{\nabla}_{\boldsymbol{\varepsilon}}$ is the $n \times 1$ vector of structural shocks given to the system in the experiment. Generally speaking, a generalized impulse response function (GIRF) estimates to the difference between the conditional expectation of the model after a shock has occurred and the conditional expectation of the same model in case no shock occurred. Clearly, $IRF_{t,H}(\mathbf{S}^t)$ may depend on the state of the system at the time of the shock. In particular, the impulse response analysis in a MSVAR framework is mainly based on Koop et al. (1996), who proposed a GIRF approach that—consistently with a Bayesian framework—considers impulse responses as random variables. In a Bayesian framework, we obtain the posterior

[13] In the homoskedastic TVP VAR case, one will simply write $\mathbf{y}_t = \mathbf{Z}_t'\boldsymbol{\beta}_t + \mathbf{A}^{-1}\boldsymbol{\Sigma}\boldsymbol{\varepsilon}_t$ and $Cov(\mathbf{u}_t) = Cov(\mathbf{A}^{-1}\boldsymbol{\Sigma}\boldsymbol{\varepsilon}_t) = \mathbf{A}^{-1}\boldsymbol{\Sigma}\boldsymbol{\Sigma}'\mathbf{A}^{-1} = \boldsymbol{\Omega}$.

distribution of the GIRFs throughout the Gibbs iterations. More precisely, once the posteriors for the VAR parameters and the structural identification matrix $\boldsymbol{\Upsilon}_t$ are available, we compute each impulse response function depending on the initial regime as well as on the future regimes $\mathbf{S}^H \equiv [S_{1t+1}\ S_{2t+1}\ ...\ S_{1t+H}\ S_{2t+H}]'$ as follows:

$$\begin{aligned} IRF_{t,0}(\mathbf{S}^t) &= \boldsymbol{\Upsilon}_t \boldsymbol{\nabla}_{\boldsymbol{\varepsilon}} \\ IRF_{t,h}(\mathbf{S}^t) &= \begin{cases} \sum_{i=0}^{h-1} \mathbf{B}_{1,S_{t+h-i}} GIRF_{t,h}(\mathbf{S}^t) & \text{if } h \geq 1 \\ 0 & \text{if } h < 0 \end{cases} . \end{aligned} \tag{21}$$

Therefore each IRF depends not only on the current state of the economy but also on the future realizations of the latent state variables. However, by integrating out the forecasted states, it is possible to obtain a different expression for the GIRF, that only requires to know the initial regime:

$$GIRF_{t,H}(\mathbf{S}^t) = \int_{S_{1t+1},S_{2t+1}} \int_{S_{1t+2},S_{2t+2}} ... \int_{S_{1t+H},S_{2t+H}} IRF_{t,h}(\mathbf{S}^t) \Pr(\mathbf{S}^H) d\mathbf{S}^H. \tag{22}$$

In practice, such multiple integration is simply performed by averaging across a large number of predicted paths for the Markov state variables.

3 Data and Summary Statistics

Our data consist of time series of ex-ante yields of several asset classes traded in the US financial market. In particular, we collect information of asset-backed securities, Treasury overnight repos, Treasury notes, investment grade and non-investment grade corporate bonds, and the stock market. We consider weekly observations from January 7, 2000 to December 27, 2013, for a total number of 5,103 observations over the nine series.

To capture the onset of the US financial crisis and an initial shock derived from the asset-backed security market, we consider both high- (i.e., Aaa rated) and low-grade (i.e., including Bbb rated) ABS indices compiled by BofA Merrill Lynch. Asset-backed securities are fixed income notes backed by cash flows from a variety of pooled receivables or loans. Generally, these pools are collateralized consumer and business loans in contrast with mortgage-backed securities, that are backed by mortgages. In particular, in the securitization process, a pooled set of assets is used by the originator to create several *tranches*, characterized by a different level of priority of claims on the collateral pool (see Agarwal, Barrett, Cun, and De Nardi, 2010, for an introduction). In case of default, the losses are absorbed by the lowest priority tranches before affecting those with higher priority. Before the subprime crisis, ABS came to represent popular investments, also because they allowed banks and other institutions subject to capital requirements to reduce the size of their balance sheets and free up capital. Over the more recent years, the investor base for ABS products has shifted away from banks and institutional investors, towards hedge funds and structured investment vehicles that profit from the difference between the yield on long-term structured products such as ABS and the short-term liabilities they issue to collect funds. Moreover, before the crisis, ABS were extensively used by financial institutions as collateral in the repo market and covered a crucial role for in their funding operations.

In the case of the repo market, we collect weekly observations on the Treasury overnight general collateral (GC) repo rate from GovPX. In a repo transaction a sale of securities is combined with an agreement to repurchase the same securities at maturity at a higher price, representing an interest rate paid to the lender. Before the subprime crisis, lenders accepted several categories of securities as collateral, including ABS. Nowadays, the U.S. repo market is dominated by transactions mainly based on US Treasury collateral. Hence our choice of the weekly yield series investigated here. The US Treasury market is instead considered in our analysis by including the yield series of 10-year constant maturity Treasury notes collected from FRED® at the Federal Reserve Bank of Saint Louis.

To have indicators of the performance of the corporate bond market, we resort to the indices computed by BofA Merrill Lynch which allow to clearly identify the maturity as well as the credit rating of each portfolio. Similarly to asset-backed security yield indices, we consider both investment- and non-investment grade bond indices and we further classify the index components according to their time to maturity: less than 3 years for the short-term bonds and more than 10 years for the long-term bonds. The non-investment grade category is represented by corporate bonds defined by a credit rating of Ccc and lower, while investment grade corporate bonds are Aa-grade and higher. Finally, we include the dividend yield rate of the S&P 500 as representative of the ex-ante yield of the US stock market. The rationale is that the empirical finance literature has repeatedly shown that the dividend-price ratio is able to predict equity returns (see Pesaran and Timmermann, 1995; Campbell and Thompson, 2008). Thus, we can reasonably compare this metric with the yield series that we use for the bond market. To address the cyclical nature of the dividend yield, we compute our weekly observations for the S&P 500 taking the moving average of the dividends paid over the previous 3 months and use it as the numerator of our ratio.[14]

Table 1 (Panel A) presents the main statistical features of our data. Figure A1 in the Appendix plots the data. Sample means and standard deviations are relatively unsurprising: ABS give yields that are slightly higher than corporate bonds, junk paper (both ABS and corporate) pays out more ex-ante than investment-grade bonds do, both ABS and corporate bonds imply higher average rates than Treasuries do, and finally the lowest yields—as one would expect—are those on repo contracts, also because these are based on very high quality collateral. Because we analyze equity dividend yields and not returns, their yield is the lowest among all security types, even though this also applies to sample standard deviation. The sample estimates of skewness and kurtosis reveal that none of the series is normally distributed; in fact, all the yields, with the only exception of Treasuries', display a positive skewness. Moreover, the ABS Aaa, repo, Treasury, and the investment grade short- and long-term corporate bond yields are characterized by a platykurtic distribution, as they imply negative

[14] A similar treatment is used by Ang and Bekaert (2007), among others.

sample values of excess kurtosis, which means that their empirical distributions display tails that are thinner than a normal. To the contrary, the sample distribution of the remaining series are leptokurtic because they imply a positive sample excess kurtosis, meaning that their tails are fatter than a normal distribution. Indeed, a standard Jarque-Bera univariate normality test rejects the null hypothesis of normality for all the series in Table 1 at standard confidence levels.

The sample correlations presented in Table 1 (Panel B), show that the bond yields included in our analysis are generally positively correlated to one another, with the exception of non-investment grade long term corporate bonds and Treasuries and repo rates which imply negative but small correlations. However, the dividend yield carries a negative correlation with all the investment-grade and all short-term fixed income rates, which is understandable in a flight-to-quality perspective, and when investors reduce the duration and hence the risk of their portfolios away from long-term towards short-term bonds during periods of financial turmoil. There is also strong evidence of a high sample correlation between the Aaa ABS and investment grade corporate bond yields, both for short and long-term maturities, which probably reflects the joint dynamics of the rating process and of pricing of the corresponding bonds during long portions of our sample leading up to the financial crisis.

In Figure 1, we plot 5-year rolling correlations between low-quality ABS yields and the yields of the other assets considered in the analysis. We observe significant time-variation in the correlations between yields immediately after the onset of the subprime crisis at the end of 2007. The figure clearly shows a strong increase in the correlations of low-quality ABS yields with high-quality ABS, corporate bonds, and stocks, going from values ranging between 0 and 0.4 to values in excess of 0.5 and close to 1 over a long spell, especially after 2011, when data from the crisis period come to dominate the 5-year rolling statistics. On the contrary, we observe a sharp decline in the correlations of low-quality ABS with repo and 10-year treasury yields, reaching values near -0.8 during 2009, and recovering to positive values during 2012. Overall, it is clear from Figure 1 that the subprime shock in the low-grade ABS market did spill over to all assets classes in the US financial market by affecting the joint dynamic of their yields. This preliminary, unconditional, evidence motivates our empirical investigation of interdependence and contagion among asset classes that follows.

4 Empirical Results

In this section, we first compare and select our preferred model from the alternatives that we have set out before and that are in principle able to capture cross-asset contagion and—in at least some cases—to disentangle contagion from simple normal interdependence between markets. In a subsequent section, we report the estimates of the time-varying coefficients for the selected model. Finally, in a third subsection, IRFs and the corresponding posterior uncertainty are estimated in order to quantify how a shock in the asset-backed security market may spill over to other markets, and when and how this may represent an instance of contagion.

4.1 Model Selection and Comparison

We start our analysis by comparing the empirical fit of the set of models proposed in Section 2. Following Koop et al. (2009), we evaluate the models by looking at the expected values of their log-likelihoods. Moreover, for mixture innovation TVP-VAR models, we use the implied, recovered information on posterior breaks probabilities to provide a convenient and intuitive measure to decide which model better fits the data. To re-cap the way we have treated them in Section 2, the set of models taken into consideration are a single-state, homoskedastic VAR which represents the building stone of the literature and yet suffers from one non-resolvable problem with the identification of contagion vs. interdependence; a homoskedastic TVP-VAR model in which contagion is detectable but may be biased by strengthened interrelationships caused by altered variances and correlations; a four-state MSVAR model with independent Markov regimes for VAR and covariance matrix coefficients, respectively; TVP-VAR with stochastic volatility, and mixture innovation TVP-VAR models (in the latter case, both with homoskedastic and heteroskedastic covariance matrices);[15] finally, a Cogley-Sargent (2005, CG) mixture innovation type model, which is a mixture model with $K_{2t} = 0$ and hence constant $\mathbf{A}$ and resulting covariances of shocks. This gives a set of eight different models.

The first column of Table 2 collects the expected values of the log-likelihood that we compute for each of the different models included in our analysis, as derived by averaging over Gibbs samples, while the remaining columns of the table report the posterior estimates of the break probabilities obtained from the mixture innovation TVP VAR models. These can be seen as posterior estimates of $\Pr(K_{it})$ with $i = 1, 2, 3$. Here the idea is that when those posterior probabilities are large and approach 1, then the mixture TVP-VAR is just mimicking the behavior of a TVP-VAR model, with or without SV; when such probabilities are smaller, then it becomes important to assess the plausibility of the dynamics they imply; if break probabilities vanish to zero, then the model is signalling that some or all the coefficients are simply constant over time. To favour comparisons, we have computed implied break probabilities (here, probabilities of switching out of a given regime, averaging over regimes using their ergodic probabilities) also from the MSVAR framework.

According to Table 2, a constant parameters, single-state homoskedastic VAR(q) is not sufficiently rich to describe the joint dynamics of the nine yield series under investigation: the corresponding expected log-likelihood (henceforth, ELL) of approximately -3,031 is considerably lower than all other estimates in the table. There is only important exception to this: a MSVAR(q) model, in which US financial markets simply shift between tranquil and crisis regimes does only slightly better, with an ELL of -2,474. In Table 2, q has always been set to 1 as VAR frameworks with longer lags never provided a better in-sample fit.[16] On the one hand, this comparison confirms the result of previous MSVAR papers and analyses (see Guo et al., 2011; Guidolin and Pedio, 2017) that capturing regime shifts may be of considerable importance. On the other hand, it is natural to ask whether and by how

[15] The difference between a full mixture innovation TVP VAR and a heteroskedastic TVP VAR is that in the former K_{1t} is estimated, in the latter $K_{1t} = 0$.

[16] Complete tabulated results are available from the authors upon request.

much even more complex VAR models with TVP or mixture innovation may improve the fit.

Table 2 implies that the benefit we obtain when moving from single-state VAR and MSVAR models to TVP-VARs is considerable. For instance, moving from a single-state VAR(1) to a comparable homoskedastic TVP-VAR(1) increases the ELL from -3,031 to -657; moving from a MSVAR(1) to a SV TVP-VAR(1) bring the ELL from -2,474 to -968. Interestingly, adding a SV component to a TVP VAR(1) reduced the in-sample fit, instead of improving it. However, this may derive from the fact that in the SV TVP-VAR framework of (10)-(11), yield volatilities are forced to break at each point in time, which may act as a confounding factor and worsen the in-sample fit provided. Crucially though, Table 2 shows that the models best fitting our data are those implying mixture innovations and hence discontinuous, infrequent breaks. For instance, a homoskedastic mixture model brings the ELL to -122 from -3,031 of the single-state VAR; a (purely) heteroskedastic bring the ELL to -178 from -968 of the SV TVP VAR models. Also in this case, the inclusion of heteroskedasticity fails to improve, per se, the fit of the model. Nonetheless the best fitting models are the complete, full mixture innovation model with ELL of -120.6 and especially the Cogley-Sargent modified mixture model with an ELL of -120.5. Even though the difference is small and possibly due to numerical approximation error (because our posteriors and ELL are simulated), there is little doubt that the improvement in the fit provided by having time-varying, infrequently breaking correlations (through $\mathbf{a}_t = \mathbf{a}_{t-1} + K_{2t}\boldsymbol{\zeta}_t$) is modest at best. This represents a first, key result of our analysis: US financial markets are characterized by breaks in the way shocks propagate over time—i.e., genuine contagion—and not simply or mostly by instability in the cross- serial correlation patterns of the shocks.

The implied break probabilities of the full and CG mixture TVP-VARs are sensible. According to the full model, breaks in VAR, correlations, and variance coefficients have smoothed probabilities of 0.110, 0.072, and 0.055, respectively. This can be read as saying that episodes of crisis and contagion on average would last 9, 14, and 18 weeks as far as spillover mechanisms, correlations, and variances are concerned. In the CG case, when the second probability is restricted to zero, the remaining two move to $\widehat{\Pr}(K_{1t}) = 0.108$ and $\widehat{\Pr}(K_{3t}) = 0.047$ and imply even longer, more plausible durations.

4.2 Patterns in Parameter Time Variation

As previously discussed, according to a few mainstream definitions (see e.g., the review by Billio and Caporin, 2010), financial contagion can be detected by investigating (formally testing) the existence of unstable parameters that control the nature and strength of the transmission mechanism of shocks over time. In particular, during contagion bouts, such parameters should move in directions that accentuate the speed and magnitude of shocks from one market to others. On the contrary, when the parameters related to speed and strength of progressive shock propagation are approximately constant over time, this should represent evidence that financial markets may still be strongly interdependent through the pairwise correlations of shocks, but will not be prone to contagion. For this reason, this section takes as given the posterior estimates obtained in Section 4.1 and proceed to examine whether the estimated (posterior means of) VAR coefficients, and volatilities/correlations of shocks have changed during the period under analysis and in what ways.[17] However, given that it is impossible to report plots for 90 VAR coefficients, we choose to report only the loadings of each yield on the low-grade ABS ones, which are the most important to feature the instance of contagion that we want to study.

4.2.1 Standard TVP VAR models

In Figure 2, we show the dynamics of the element of $\boldsymbol{\beta}_t$ that loads on past values of ABS yields of the lowest credit quality in our data set for a standard, single-state VAR, a homoskedastic TVP-VAR, and SV TVP-VAR. The striking result is that a constant-coefficient, standard VAR may completely hide strong cross-market linkages that develop during times of crisis. With the only exception of investment-grade long-term corporate yields (and of course of low-quality ABS rates on their own past), in all other plots the constant VAR coefficient is very close to zero and always fails to be significant using 90 percent confidence regions (this is checked in an unreported outputs). However, all risky assets are characterized by a spike of their loadings on past low-grade ABS yields in correspondence to the 2008-2009 crisis. In some cases—notably, investment-grade short-term, non-investment-grade long-term corporate rates—the spike is visible and brings the VAR coefficient under discussion from small and non-positive values to high and positive ones, implying that jumps in ABS yields do get transmitted to these markets in sizeable and "abnormally high" ways that clearly define a contagion episode. The effect is more muted, but still present, in the case of investment-grade short-term corporate and equity dividend yields.[18] Interestingly, these contagion effects show up irrespective of whether stochastic volatility is modelled or not, although the spikes are sharper when the TVP-VAR model is homoskedastic. By the end of our sample, most values of the posterior mean coefficients have settled back to their pre-crisis levels, which may be taken as an indication that the stress conditions recorded between 2008 and 2009 had eventually vanished by the end of 2013.

Figure 3 performs a similar comparison with reference to the posterior estimates of the own-shock volatility parameters for the three models already covered by Figure 2. Obviously, only one such estimate is time-varying and corresponds to the SV framework. Two effects are notable. First, a single-state, constant parameter VAR implies estimates of the standard deviation of the shocks that are always massively larger than under a homoskedastic TVP-VAR; in both cases, the estimated coefficients are constant over time, but its level is completely different, easily reaching peaks in which if the TVP aspect is ignored, the posterior estimate of volatility is between 2 and 4 times larger (it is almost 16 times larger in the case of non-investment grade, short-term yields). This illustrates

[17] Due to space constraints, in this and the following subsections, we limit our analysis to the models with the highest empirical ELLs. Complete tabulated and plotted estimates and analysis for the unreported models are available from the authors upon request. Following Primiceri (2005), we use the first 40 observations (from January to September 2000) as a training sample to define empirical prior distributions. Thus, our estimates refer to a period slightly shorter than the original dataset (i.e., from October 2000 to December 2013).

[18] In the case of repo and 10-year Treasury yields, the increase in their loadings on past low-grade ABS shocks occurs but in the sense that the coefficient climbs from large negative values towards zero. This means that while before the financial crisis there had been a visible flight-to-quality effect, this weakens between 2008 and 2009.

how dangerous can it be to ignore the strong evidence (see Table 2) of time-varying conditional mean coefficients in the data: most randomness in yields would get incorrectly imputed to unpredictable shocks, while we know that most of it (typically between half and three-quarters of it) actually derives from variation in VAR coefficients. Second, when variability in the posterior mean of the standard deviation of shocks is allowed, Figure 3 shows that all series (but the dividend and non-investment grade short-term yields) repeatedly spike up between 2007 and 2009.[19] Of course, this is relatively unsurprising because the crisis was a jittery period of market turmoil. However, in our case, this finding is important for another reason: a TVP-VAR model with stochastic volatility can then tell apart the effects of contagion from the instability that affects the scale of the shocks. This is very important, because it guarantees that the result in Figure 2—that low-quality ABS shocks get transmitted with a different, stronger intensity during the financial crisis—holds net of the fact that shocks carry a larger magnitude during the same period. This point is further emphasized by Figure 4, which shows (for all models, but here our attention goes to the SV TVP-VAR case) that all pairwise correlations between shocks to each of the eight markets and low-grade ABS yields are in generally time-varying, but more importantly *decline* during the financial crisis; in words, the glamorous, strong co-movements (one could argue, co-jumps) during the Great Financial Crisis were purely induced by contagion and in no plausible way by simple increased interdependence.[20]

4.2.2 Mixture Innovation TVP VAR models

Figure 5 reports plots comparable to Figure 2, but referring now to the four types of mixture TVP-VAR models entertained in this paper (i.e., full model with no restrictions on $\mathbf{K}_t$, homoskedastic with $K_{2t} = K_{3t} = 0$, heteroskedastic with $K_{1t} = 0$, and the Cogley-Sargent variant with $K_{2t} = 0$). As in Figure 2, also in Figure 5 we focus on the VAR coefficient loading yields on past values of low-quality ABS yields, even though complete (but copious) results are available upon request. On the one hand, the results in Figure 5 are less homogeneous and hence impressive when compared to the simpler, TVP-VAR cases. However, also in this case we obtain ample evidence that VAR coefficients—in the three models in which they are allowed to vary over time—do tend to move considerably in two periods: during the "dot-com" equity market bust of 2000-2002 and during the 2008-2009 financial crisis.[21] Yet, the way the posterior means of the coefficients move is now more complex, but at the same time also easier to interpret: all risky asset classes (i.e., all ABS, investment-grade corporate, junk long-term corporate, and equities) but one (non-investment-grade short-term corporate) record an upward jump to positive and relatively high posterior mean estimates of the values of their exposure to low-quality ABS shocks between 2008 and 2009; relatively riskless assets (Treasuries and especially repo contracts)

[19] The volatility of dividend yield shocks is indeed time-varying but the spikes occur in 2002-2003. In the case of junk short-term corporate rates, their volatility is essentially flat because our estimates reveal that during 2008-2009 is their connection to other asset markets that undergoes a deep change.

[20] Interestingly, the strength of the correlation between yields are not affected by whether one estimates constant correlations under a standard VAR(1) or in a homoskedastic TVP VAR model.

[21] Also in Figure 5, a purely heteroskedastic mixture VAR model delivers posterior mean of the loadings on low-quality ABS shocks that are negligible, which is of course rather hard to make sense of. As we have seen in Table 2, this model is dominated by other mixture models, in which a TVP VAR component is required.

imply instead a muted or even negative reaction, i.e., their exposure to ABS, subprime-type shocks declines during the financial crisis. Such a contemporaneous contagion and flight-to-quality effects are exactly what one would expect in the light of the now ponderous literature on the crisis. However, while in the case of a few markets, the size of the contagion estimated from the different models falls in a very narrow range, in a few of the plots in Figure 5 (Aaa ABS, equities, and investment-grade corporate yields), more dispersion emerges. Also this effect is sensible, as these high-quality, triple A-rated fixed income securities are probably impacted by low-quality ABS shocks to a smaller extent and our models do capture that also through a bigger dispersion of the measurable, posterior effects. The mystery in Figure 5 is posed by non-investment grade, short term corporate yields, that react to the financial crisis with their loadings on low-grade ABS yield shocks declining to even more negative values, almost as if in the face of bearish sub-prime type ABS markets, investors would react by moving their funds towards short-term junk bonds.[22]

It is of course important to also ask whether the dynamics in the loadings/VAR coefficients in Figure 5 arise from sensible ex-post inferences on the location in time of breaks (as governed by K_{1t}) and their posterior, smoothed probabilities. Figure 6, top panel, shows the smoothed series $\widehat{\Pr}_{t|T}(K_{1t})$. Impressively, the posterior of $\widehat{\Pr}_{t|T}(K_{1t})$ hardly depends on the specific model estimated, when K_{1t} is not restricted to always take a value of zero. This is a sign of robustness of our results because in no ways such similarities have been imposed while our priors are generally weak. Breaks in VAR coefficients are frequent, but they tend to concentrate in two specific periods, as one would expect: 2000-2001 and then late 2008 and 2009; however, also 2006 is characterized by some evidence of shifting conditional mean coefficients. In fact, over the 2000-2001 sample there is an ex-post probability of 0.26 of recording a conditional mean parameter break on every week; over the 2008-2009 sub-sample, such a probability is estimated at 0.16; over the remaining, non-crisis sample, the posterior frequency is instead slightly in excess of 0.02, as one would expect. This establishes that according to mixture TVP-VAR models, it is crisis samples that are dominated by breaks in coefficients.

Figure 7 performs an operation similar to Figure 3 with reference to time-varying volatility coefficient posterior estimates and delivers a qualitatively similar point: in this case, and according to all models that do not restrict K_{3t} to be zero (homoskedastic mixture TVP-VAR), we record a slight increase in the variance of shocks in 2008-2009, followed by a series of (relatively sparse) breaks that cause posterior volatility to progressively decline over time (this effect is weaker for stocks). Even though the exact levels of the inferred volatility of shocks differ across models, the extent and timing of the upsurge between 2008 and 2009 followed by a steep reduction are similar. The decline in the recorded volatility of fixed income yields is of course one of the (possibly intended) effects of the unconventional monetary policies pursued by the Federal Reserve between 2009 and 2013 (see, e.g., Yang and Zhou, 2016). Figure 6, middle panel, reflects (without recording the signs of the changes in volatility of

[22] In contrast with the traditional literature that defines contagion as an increase in cross-markets linkages, Corsetti, Pericoli, and Sbracia (2011) show that, depending on the nature of the crisis, the related parameter values could either decrease or increase during periods of turbulence in the markets. In line with their findings, our cross-assets relationships do not always strengthen in distressed periods. However, we remain surprised by our empirical finding.

course) these patterns and shows a dense barrage of frequent, almost weekly adjustments driven by $K_{3t} = 1$ between late 2008 and mid-2010, followed by a more tranquil period. Interestingly, in all models, between 2003 and early 2007, $\widehat{\Pr}_{t|T}(K_{3t}) = 0$.

Figure 8 reports similar findings with reference to correlation coefficients concerning shocks to each of the asset classes and low-grade ABS yields.[23] For most models and markets, such correlations appear to be declining—a picture of a US financial markets in which spillover across markets are getting stronger and are subject to actual contagion effects, but in which unanticipated shocks become less interrelated. Even though there is pale evidence of an increase in shock correlations in correspondence to the US financial crisis, this effect is weak and appears to strictly depend on the model investigated. In fact, Figure 6 shows that the smoothed $\widehat{\Pr}_{t|T}(K_{2t})$ are generally small, with a minor uptick in 2008-2009 only under a very specific model under the many frameworks considered. All in all, Figure 5 and 8 jointly considered show that during the financial crisis (and possibly, also during the "dot-com" market bust in 2000-2001) there was a contagion of considerable strength from the low-quality ABS market to other risky segments of the US fixed income markets, that hardly depends on the specific model considered and when any heteroskedastic effect is netted out by performing posterior estimation in heteroskedastic models, as discussed by Koop et al. (2009). On the contrary, there is no strong evidence of an increase strength of the interdependence among markets, i.e., of the fact that correlated shocks hit all the markets simultaneously.

4.3 Impulse Response Functions

Besides disentangling contagion from interdependence, the ultimate objective of our analysis is to identify the channels through which negative shocks in one market propagate to others. For this reason, we present and compare the main features of the IRFs we have estimated under the different models presented in Section 2. In particular, our goal is to study how a negative shock affecting the lower-grade ABS market in 2007 may have been transmitted to the other asset classes investigated in our paper. Thus, in what follows we simulate a one-standard deviation shock to the low-grade ABS yield and recursively trace its effects on the remaining markets over a period of 21 weeks.[24]

4.3.1 Standard TVP VAR models

The time-varying nature of the parameters of all these models implies that also the implied IRFs change at each point in our sample, as the smoothed parameter values are updated in the process of applying the Gibbs sampler. In order to get a complete representation of our results, Figures 8 through 10 provide both three-dimensional and a bi-dimensional plots of the IRFs from two different

[23] Correlations are time-varying also in the C-S TVP VAR framework because, even though $K_{2t} = 0$, K_{3t} remains time-varying and—for given covariance—the vary variation of the standard deviations is sufficient to induce time variation in correlations.

[24] We omit confidence intervals from plots of IRF plots to avoid making them overcrowded. Of course, confidence intervals are important for inferential purposes, but considering the high number of models under analysis, and for clarity in their comparison, our priority is to showcase the robustness of our findings concerning contagion and interdependence across a set of highly flexible models.

TVP-VAR models. Figures 8 and 10 show the overall evolution of the IRFs between October 2000 to December 2013 for four asset classes that are representative of possible patterns (a complete set of 9 plots appears in Figures A2 and A4 in the Appendix); Figure 10 (an equivalent set of plots under SV is reported in an Appendix) are instead based on four selected dates in order to obtain a more intuitive comparison between different periods of interest. mid-October 2000 and mid-September 2008 capture the most acute effects of the dot-com crisis and of the sub-prime crisis, respectively; early January 2005 and the last week of December 2013 describe how the transmission mechanisms settled down some years after the end of each of the two crises. To keep the plots as readable as possible, we elect not to report credibility regions/intervals of our posterior IRF estimates; these are available upon request, but a Reader must be warned that these are generally quite wide, because of the high number of parameters that characterize this class of models.

Figures 8 and 9 refer to the homoskedastic version of the TVP-VAR model, while the results for the TVP-VAR with stochastic volatility are presented in Figure 11 (and A3 in the Appendix). According to our estimates, the two models do not significantly differ in terms of the implied IRFs, as they capture similar contagion effects: most IRFs seem to be relatively flat and quickly converge—when reading the plots right to left, from when the shock is simulated to 21 weeks later—to zero. However, all of them show a rather different behavior in correspondence to the weeks between late 2008 and mid-2009: especially for the riskier assets, the IRFs are much larger and most importantly the systematically increase over our 21-week horizon (however, they stabilize and eventually decline, as they should given that stationarity of the models has been enforced during the estimation process).[25] Crucially, all IRFs jump up and imply much stronger effects (between 4 and 20 times the impact recorded outside crisis periods), which we impute to the existence of contagion effects. Comparing Figure 9 and 11 reveals that, by including stochastic volatility, TVP-VAR models improve their ability of detecting episodes of contagion, in the sense that the breaks altering the position and shape of the IRFs become stunningly large in Figure 11, when one nets out the effects produced by heteroskedasticity. This result is consistent with Primiceri (2005) who discusses how modelling stochastic volatility is necessary to fully capture the dynamics of contagion.

4.3.2 Mixture Innovation TVP VAR models

Due to their similarities, we present the results we have obtained from the mixture innovation TVP-VAR models only with reference to the full model. Results for the CG mixture framework were visibly indistinguishable for our purposes. Figures 11 and 12 show the three-dimensional and the bi-dimensional plots, respectively. In general, despite the introduction of the mixture innovation variables, the overall pattern of the IRFs is similar to the ones we have reported for the simpler TVP-VAR models in Figures 8-10. However, there are few exceptions. In particular, comparing

[25] Also in Figures 8-10, the effect of a low-grade ABS shock is negative on repo and Treasury yields. It is interesting to note that the dot-com crisis caused very different reactions when compared to the recent crisis. In particular, during the dot-com bust, 10-year Treasury rates display a non-monotonic, positive response, the opposite of the negative, immediate reaction that characterizes the subprime crisis.

Figures 11 and 10, we can see that, for most yield series, the IRFs in the mixture models behave less erratically and present less abrupt shifts than in the standard TVP-VAR case, with or without stochastic volatility. Especially during the two periods of crisis and contagion that we have identified, a shock to lower-grade ABS yields generates weaker and smoother reactions in other series, especially in corporate rates, vis-a-vis the case in which VAR parameters are forced to undergo continuous changes in Figure 11. These results are similar the ones discussed in Section 4.1, where contagion was starker when time-variation in parameters was assumed to be pervasive.

5 Further Discussion: Transmission Channels

In this Section, we explain how the results presented in Section 4—where we have established that contagion spillovers were at work over and on top of (otherwise weak) interdependence effects—can be used to detect which contagion channels were operating in the period under investigation. In practice, we start by defining the most sizeable, relevant effects that arise in each market when each of these contagion channels is at work, and then use these indicators to establish which of the transmission mechanisms drove the effects observed during the financial crisis. The analysis is carried out for the CG version of the mixture innovation TVP-VAR, but results are not sensitive to adopting other frameworks of analysis that include mixture components.

5.1 Flight-to-liquidity channel

According to the flight-to-liquidity channel, a shock to one market causes an increase in investor's demand for highly liquid assets, and the consequent appreciation of these assets. On the contrary, the price of the other, non-liquid assets declines. Thus, the effects of this channel could be easily detected by studying the behavior (e.g., the IRFs) of Treasury yields, certainly the most liquid asset in the menu we have studied.[26] Moreover, the demand for Treasuries is also reflected by the reaction of the Treasury repo rate. As noted by Banerjee and Graveline (2013), repo transactions could be either cash-driven or security-driven. When the second situation occurs, investors enter into repo transactions with the objective of borrowing Treasury bonds, and thus, are likely to provide cash at more favorable conditions (i.e., lower repo rates). For this reason, we may refer to the IRFs of Treasury as well as the overnight repo rates, to investigate the presence of flight-to-liquidity phenomena.

Our earlier results show that Treasuries negatively react to a shock to low-quality ABS during all the periods under analysis, as one would expect under the flight-to liquidity channel; the only exception is the period 2003-2006, when the corresponding IRF turned positive. This is, during expansions/bull markets, when the demand of Treasury bonds declines as investors shift their investment preferences towards more profitable assets, reaching for yield. These conclusions are also confirmed by the IRFs that we obtained for the repo market, which are negatively affected by a shock to low-grade ABS yields

[26] Also stocks are liquid, but here we are studying a large index that also contains stocks that are surely less liquid than 10-year Treasuries.

throughout the entire sample. Interestingly, this effect strongly intensified from the outbreak of the subprime crisis till the end of 2009, when repo rates required by lenders steeply decreased. It is likely that investors entered these transactions with the goal of borrowing Treasuries used as collateral.

5.2 Risk premium and flight-to-quality channels

The risk premium channel implies that, after a shock to one market, the overall risk aversion of the average investor increases, leading to higher risk premia across the board and hence higher required yields. Similarly, under the flight-to-quality channel, investors react to a shock by selling risky assets and simultaneously purchasing assets that are perceived to be safer because they have a stronger credit quality (probability of repayment). The effect is that the risk premium on the former assets is expected to climb, while the opposite occurs to the safest assets (see e.g., Gonzalo and Olmo, 2005). For instance, Pasquariello (2014) shows that investors demanded significantly higher risk premia to hold stock and currency portfolios during the 2008 great financial crisis in the US. This increase in risk premia is explained by what he defines financial market *dislocations*, meaning, large, widespread asset mispricings in periods of distress.

We indirectly investigate the presence of these two channels by looking at the contemporaneous effects of a shock on the yield series of the riskiest among the assets and at the relatively safe asset, proxied by the repo market yield. Our results show that the repo rate presents negative IRFs throughout the entire sample period. Thus, we dismiss the possibility that the positive risky yield IRFs were due to higher risk-free rates and consequently interpret such results as signals of an increase in their risk premia. Based on this finding, we study the evolution of the transmission of the shocks to low-grade ABS to investment- and non-investment grade corporate bonds, and the dividend yield, to detect signs of risk premium and flight-to quality channels. In Section 4 we have reported a simultaneous depreciation of all these risky yield series and this supports the presence of a risk premium channel, while the combination of positive effects on the riskier markets (i.e., non-investment grade corporates and the dividend yield) and contemporaneous negative effects on the safer markets (i.e., investment-grade corporates) points towards a flight-to-quality channel.[27]

Both short- and long-term investment-grade corporate bonds display a strong positive response during the subprime crisis. However, in Section 4 a difference emerged that is worth mentioning. Short term corporates are characterized by a response that gradually moved from negative to positive values during the crisis, whereas for long-term corporates, the change is much more abrupt, with a rapid decline to negative values in 2007, suddenly followed by a positive reaction in 2008. Except for this short-lived negative response, our results for the safer corporate bond market support the hypothesis that,

[27] The ABS higher grade market reacts slightly negatively to a shock to the ABS yields in most periods and especially in 2002. On the opposite, however, the effect becomes temporarily positive around September 2008, in correspondence to Lehman's demise. Thus, at the beginning of the crisis, when the distress was still restricted to low-quality ABS, the high-grade ABS yield decreased, supporting the presence of flight to quality channels. As the shock spread over the entire financial system, however, in Section 4 we have recorded a depreciation of high-quality ABS, which can be explained by a subsequent positive effect on the correspondent risk premium. This reaction is not surprising, since the two ABS markets are strongly linked because of their nature and investors' demand for the overall class of asset-backed securities likely decreased when the crisis spread further.

during the recent financial crisis, a risk premium channel has been at work. These findings could be explained by the extremely widespread uncertainty that characterize the US financial markets during the crisis and are consistent with the results of Gorton (2009), who claims that, during the subprime crisis, the depreciation of the corporate bond market was driven by generalized, high risk aversion. More precisely, as the ABS market collapsed, it became extremely difficult for financial institutions to obtain collateralized loans through repo transactions, because of the stringent conditions imposed by lenders. Due to their high credit ratings, investment grade corporate bonds provided an attractive alternative financing strategy as they could be sold in the market to obtain immediate funding.

The IRFs of non-investment grade corporate bonds reveal instead heterogeneous reactions depending on whether we consider short- or long-term paper: yields on short-term corporates remarkably decrease while the opposite occurs to long-term ones. These results suggest that risk premium channels affected only long-term, risky corporate paper. An explanation of this finding is that, after the negative shock to the ABS Aa-Bbb, the worst companies included in the non-investment corporate bond portfolio immediately defaulted and the credit level of the overall class increased. The effects of this improvement are reflected by the appreciation of short- term corporate bonds which are usually issued by companies riskier than the ones included in the long-term portfolio (see Helwege and Turner, 1999).

5.3 Correlated information channel

The correlated information channel predicates that a negative shock in one market generates widespread but immediate effects in all markets for which the shock carries payoff-relevant information. This reaction is due to the fact that investors rapidly incorporate the news from one market in the prices of other assets. Because of the abrupt nature of this mechanism, we associate the correlated information channel to the dynamic and non-linear contagion effect captured by our time-varying parameters model. In particular, we look at the evolution of the difference between the IRFs computed under a mixture innovation TVP-VAR model and a simple, single-state VAR model with constant parameters. For each variable, we analyze the reaction recorded one period after the initial shock, since we are only interested in immediate effects.

Our results in Section 4 show that the effects of the correlated information channel get stronger at the beginning of our sample (2000-2001) and during the financial crisis (from mid-2007 to mid-2009). In particular, during the subprime crisis, part of the positive reactions (3 bps in just a week) can be imputed to the correlated information channel. Moreover, the negative reaction registered in the repo and Treasury markets appears to be affected by this channel by roughly 3 bps. In a similar way, the IRFs computed for investment- and non-investment grade long-term corporate bonds as well as equities are only slightly affected by the correlated information channel (4 bps). On the opposite, we found significant evidence of information-driven effects for short-term corporates. Indeed, the huge yield increase characterizing the former (150 bps) as well as the abrupt decline experienced by the latter (130 bps) can be almost totally explained by a correlated information channel.

6 Conclusions

We investigate the occurrence of cross-asset contagion by modelling time-varying conditional mean parameters and heteroskedasticity in US ex-ante yields. We study whether and when the transmission mechanisms of shocks across different markets have changed over the period under analysis (2000-2013) and identify the nature of such changes. We exploit the negative shock observed in the ABS Aa-Bbb market at the onset of the subprime crisis to analyze how this shock propagates across different markets, especially to high-grade ABS, Treasury, corporate, and equity yields. Importantly, we depart from the recent literature (see, e.g., Longstaff, 2010; Guo, et al., 2011; Guidolin and Pedio, 2017) by using a Bayesian approach that help us to handle instabilities in the estimated models and by experimenting with a more flexible mixture-innovation VAR model with time-varying parameters in which the timing as well as the nature of the breaks are directly determined by the data.

We find that cross-assets relationships, as well as the volatility of the exogenous shocks in our model are all changing over time, but with different timings. The posterior estimates of break probabilities that we have obtained under the mixture innovation TVP-VAR model, suggest that shifts in conditional mean coefficients intensely occurred at the end of the dot-com bubble and at the outset of the subprime crisis. On the opposite, breaks in the volatility of shocks are only identified after 2009. These results are in line with Primiceri (2005) who supports the crucial importance of jointly capturing time variation both in conditional mean and in covariance matrix coefficients.

Our results provide evidence of cross-asset contagion effects during the financial crisis: our GIRF analysis suggests that the transmission mechanism of shocks has repeatedly changed over the period under study, revealing sustained interconnections pointing to contagion effects after September 2008, when the US financial system entered in the crisis. According to the results obtained from a rich mixture innovation model, shifts in the transmission mechanisms were mainly driven by time-varying VAR coefficients that significantly changed during the 2008-2009 turmoil. A detailed analysis of estimated IRFs has also revealed that the flight-to-liquidity as well as the risk premium and correlated information channels were at work when the subprime crisis spread over the US financial system.

References

[1] Abbate, A., Eickmeier, S., Lemke, W., & Marcellino, M. (2016). The changing international transmission of financial shocks: evidence from a classical time-varying FAVAR. *Journal of Money, Credit and Banking*, 48(4), 573-601.

[2] Ang, A., & Bekaert, G. (2007). Stock return predictability: Is it there? *Review of Financial Studies*, 20(3), 651-707.

[3] Arakelian, V., & Dellaportas, P. (2012). Contagion determination via copula and volatility threshold models. *Quantitative Finance*, 12(2), 295-310.

[4] Bacchiocchi, E. (2017). On the identification of interdependence and contagion of financial crises. *Oxford Bulletin of Economics and Statistics*, 79(6), 1148-1175.

[5] Bai, J., Julliard, C., & Yuan, K. (2012). Eurozone sovereign bond crisis: Liquidity or fundamental contagion. Working paper, Federal Reserve Bank of New York.

[6] Banerjee, S., & Graveline, J. J. (2013). The cost of short-selling liquid securities. *Journal of Finance*, 68(2), 637-664.

[7] Billio, M., & Caporin, M. (2010). Market linkages, variance spillovers, and correlation stability: Empirical evidence of financial contagion. *Computational Statistics & Data Analysis*, 54(11), 2443-2458.

[8] Brenner, M., Pasquariello, P., & Subrahmanyam, M. (2009). On the volatility and comovement of US financial markets around macroeconomic news announcements. *Journal of Financial and Quantitative Analysis*, 44(6), 1265-1289.

[9] Brunnermeier, M. K. (2009). Deciphering the liquidity and credit crunch 2007–2008. *Journal of Economic Perspectives*, 23(1), 77-100.

[10] Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies*, 21(4), 1509-1531.

[11] Carter, C. K., & Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 541-553.

[12] Casarin, R., Tronzano, M., & Sartore, D. (2015). Sovereign risk and contagion effects in the Eurozone: A Bayesian stochastic correlation model. In *Advances in Statistical Models for Data Analysis*, pp. 27-34. Springer.

[13] Casarin, R., Sartore, D., & Tronzano, M. (2016). A Bayesian Markov-switching correlation model for contagion analysis on exchange rate markets. *Journal of Business & Economic Statistics*, forthcoming.

[14] Ciccarelli, M., & Rebucci, A. (2007). Measuring contagion and interdependence with a Bayesian time-varying coefficient model: An application to the Chilean FX market during the Argentine crisis. *Journal of Financial Econometrics*, 5(2), 285-320.

[15] Ciccarelli, M., & Rebucci, A. (2003). Measuring contagion with a Bayesian, time-varying coefficient model. Working paper No. 263, European Central Bank.

[16] Cogley, T., & Sargent, T. J. (2005). Drifts and volatilities: monetary policies and outcomes in the post WWII US. *Review of Economic Dynamics*, 8(2), 262-302.

[17] Connolly, R. A., Stivers, C., & Sun, L. (2007). Commonality in the time-variation of stock-stock and stock-bond return comovements. *Journal of Financial Markets*, 10(2), 192-218.

[18] Corsetti, G., Pericoli, M., & Sbracia, M. (2011). Correlation analysis of financial contagion. In *Financial Contagion: the Viral Threat to the Wealth of Nations*, pp. 11-33. Wiley.

[19] Dornbusch, R., Park, Y. C., & Claessens, S. (2000). Contagion: understanding how it spreads. *The World Bank Research Observer*, 15(2), 177-197.

[20] Flavin, T. J., & Sheenan, L. (2015). The role of US subprime mortgage-backed assets in propagating the crisis: contagion or interdependence? *North American Journal of Economics and Finance*, 34, 167-186.

[21] Forbes, K. J., & Rigobon, R. (2002). No contagion, only interdependence: measuring stock market comovements. *Journal of Finance*, 57(5), 2223-2261.

[22] Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models.* Springer Science & Business Media.

[23] Dungey, M., Fry, R., Gonzalez-Hermosillo, B. & Martin, V. (2007). Contagion in international bond markets during the Russian and the LTCM crises. *Journal of Financial Stability*, 2(1), 1-27.

[24] Eickmeier, S., Lemke, W., & Marcellino, M. (2015). Classical time varying factor-augmented vector auto-regressive models—estimation, forecasting and structural analysis. *Journal of Royal Statistical Society–Series A*, 178(3), 493-533.

[25] Gerlach, R., Carter, C., & Kohn, R. (2000). Efficient Bayesian inference for dynamic mixture models. *Journal of the American Statistical Association*, 95(451), 819-828.

[26] Giordani, P., & Kohn, R. (2008). Efficient Bayesian inference for multiple change-point and mixture innovation models. *Journal of Business & Economic Statistics*, 26(1), 66-77.

[27] Gonzalo, J. & Olmo, J. & (2005). Contagion versus flight to quality in financial markets. Working paper, Universidad Carlos III Madrid.

[28] Gorton, G. B. (2009). Information, liquidity, and the (ongoing) panic of 2007. Working paper No. 14649, National Bureau of Economic Research.

[29] Guidolin, M., & Pedio, M. (2017). Identifying and measuring the contagion channels at work in the European financial crises. *Journal of International Financial Markets, Institutions and Money*, forthcoming.

[30] Guo, F., Chen, C. R., & Huang, Y. S. (2011). Markets contagion during financial crisis: A regime-switching approach. *International Review of Economics & Finance*, 20(1), 95-109.

[31] Harrison, J., & West, M. (1997). *Bayesian Forecasting and Dynamic Models.* New York: Springer.

[32] Helwege, J., & Turner, C. M. (1999). The slope of the credit yield curve for speculative-grade issuers. *Journal of Finance*, 54(5), 1869-1884.

[33] Jawadi, F., Louhichi, W., & Cheffou, A. I. (2015). Testing and modeling jump contagion across international stock markets: A nonparametric intraday approach. *Journal of Financial Markets*, 26, 64-84.

[34] Kaabia, O., Abid, I., & Guesmi, K. (2013). Does Bayesian shrinkage help to better reflect what happened during the subprime crisis? *Economic Modelling*, 31, 423-432.

[35] Kaminsky, G. L., Reinhart, C. M., & Vegh, C. A. (2003). The unholy trinity of financial contagion. *Journal of Economic Perspectives*, 17(4), 51-74.

[36] Kim, C. J., & Nelson, C. R. (1999). *State-Space models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications.* MIT Press.

[37] King, M. A., & Wadhwani, S. (1990). Transmission of volatility between stock markets. *Review of Financial Studies*, 3(1), 5-33.

[38] Kodres, L. E., & Pritsker, M. (2002). A rational expectations model of financial contagion. *Journal of Finance*, 57(2), 769-799.

[39] Kohonen, A. (2014). Transmission of government default risk in the eurozone. *Journal of International Money and Finance*, 47, 71-85.

[40] Koop, G., & Korobilis, D. (2010). Bayesian multivariate time series methods for empirical macroeconomics. *Foundations and Trends in Econometrics*, 3(4), 267-358.

[41] Koop, G., Leon-Gonzalez, R., & Strachan, R. W. (2009). On the evolution of the monetary policy transmission mechanism. *Journal of Economic Dynamics and Control*, 33(4), 997-1017.

[42] Koop, G., Pesaran, M. H., & Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics*, 74(1), 119-147.

[43] Krishnamurthy, A., & Vissing-Jorgensen, A. (2012). The aggregate demand for Treasury debt. *Journal of Political Economy*, 120(2), 233-267.

[44] Longstaff, F. A. (2010). The subprime credit crisis and contagion in financial markets. *Journal of Financial Economics*, 97(3), 436-450.

[45] Marron, J., & Wand, M. (1992). Exact Mean Integrated Squared Error. *The Annals of Statistics*, 20(2), 712-736

[46] Mumtaz, H., & Surico, P. (2009). The transmission of international shocks: a factor-augmented VAR approach. *Journal of Money, Credit and Banking*, 41(1), 71-100.

[47] Pasquariello, P. (2014). Financial market dislocations. The Review of Financial Studies, 27(6), 1868-1914.

[48] Pericoli, M., & Sbracia, M. (2003). A primer on financial contagion. *Journal of Economic Surveys*, 17(4), 571-608.

[49] Pesaran, M. H., & Pick, A. (2007). Econometric issues in the analysis of contagion. *Journal of Economic Dynamics and Control*, 31(4), 1245-1277.

[50] Pesaran, M. H., & Timmermann, A. (1995). Predictability of stock returns: robustness and economic significance. *Journal of Finance*, 50(4), 1201-1228.

[51] Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. The *Review of Economic Studies*, 72(3), 821-852.

[52] Sensier, M., & Van Dijk, D. (2004). Testing for volatility changes in US macroeconomic time series. *Review of Economics and Statistics*, 86(3), 833-839.

[53] Stock, J. H., & Watson, M. W. (1996). Evidence on structural instability in macroeconomic time series relations. *Journal of Business & Economic Statistics*, 14(1), 11-30.

[54] Yang, Z., & Zhou, Y. (2016). Quantitative easing and volatility spillovers across countries and asset classes. *Management Science*, 63(2), 333-354.

Table 1: Summary Statistics

Panel A – Descriptive statistics for weekly observations on (annualized percentage) yields on ABS, Treasury, corporate bond and stock yields for the sample January 7, 2000 - December 27, 2013.

	Mean	Std. Deviation	Kurtosis	Skewness	Jarque-Bera	Min	Max
ABS AAA	3.65	2.06	-1.04	0.20	38.00	0.68	8.58
ABS AA-BBB	6.64	3.50	4.91	2.01	1207.08	2.21	21.04
Repo rate	2.12	2.08	-0.95	0.70	94.85	-0.01	6.64
Treasuries	3.91	1.16	-0.44	-0.19	10.50	1.47	6.77
Investment grade ST	3.60	1.94	-0.88	0.34	38.06	0.82	8.78
Investment grade LT	6.43	1.03	-0.46	0.19	10.78	4.32	9.27
Non-Investment grade ST	20.40	9.26	3.59	1.75	754.49	10.37	68.76
Non-Investment grade LT	12.78	4.77	7.46	2.59	2472.56	8.25	38.49
Dividend Yield	1.92	0.45	1.74	0.67	144.07	1.02	3.94

Panel B – Sample correlations

	ABS AAA	ABS AA-BBB	Repo rate	Treasuries	Inv. grade ST	Inv. grade LT	Non-Inv. grade ST	Non-Inv. grade LT	Dividend Yield
ABS AAA	1.000								
ABS AA-BBB	0.646	1.000							
Repo rate	0.787	0.119	1.000						
Treasuries	0.768	0.237	0.799	1.000					
Investment grade ST	0.984	0.617	0.801	0.755	1.000				
Investment grade LT	0.822	0.723	0.485	0.733	0.807	1.000			
Non-Investment grade ST	0.588	0.608	0.240	0.328	0.586	0.755	1.000		
Non-Investment grade LT	0.381	0.729	-0.045	-0.043	0.398	0.529	0.772	1.000	
Dividend Yield	-0.296	0.238	-0.575	-0.695	-0.295	-0.344	-0.143	0.300	1.000

Table 2: Model Selection and Comparison

The table shows the empirical estimates of the (expected, across Gibbs simulations) log-likelihood and of the probabilities of breaks occurring for a range of models. Pr(K_{1t} =1) is the posterior mean frequency of a break in the VAR coefficients (i = 1), correlations (i = 2), and variances (i = 3). TVP means Time-Varying Parameters. VAR means Vector Autoregressive. MS means Markov switching. MI means mixture innovations. CG means Cogley-Sargent, and represents the special case of a MI TVP VAR model under the restriction K_{2t} = 0.

	Expected Log Likelihood	Pr(K_{1t} =1)	Pr(K_{2t} =1)	Pr(K_{3t} =1)
Single-state VAR	-3,031.0	-	-	-
MSVAR	-2,473.8	0.234	0.175	0.175
TVP VAR Homoskedastic	-657.0	1	0	0
TVP VAR with Stochastic Volatility	-968.4	1	1	1
Full MI TVP VAR	-120.6	0.110	0.072	0.055
MI TVP VAR Homoskedastic	-121.7	0.109	-	-
MI TVP VAR Heteroskedastic	-178.1	0.000	0.069	0.036
CG-MI TVP VAR	-120.5	0.108	-	0.047

Figure 1: Rolling Window Correlations with Sub-Prime ABS Yields

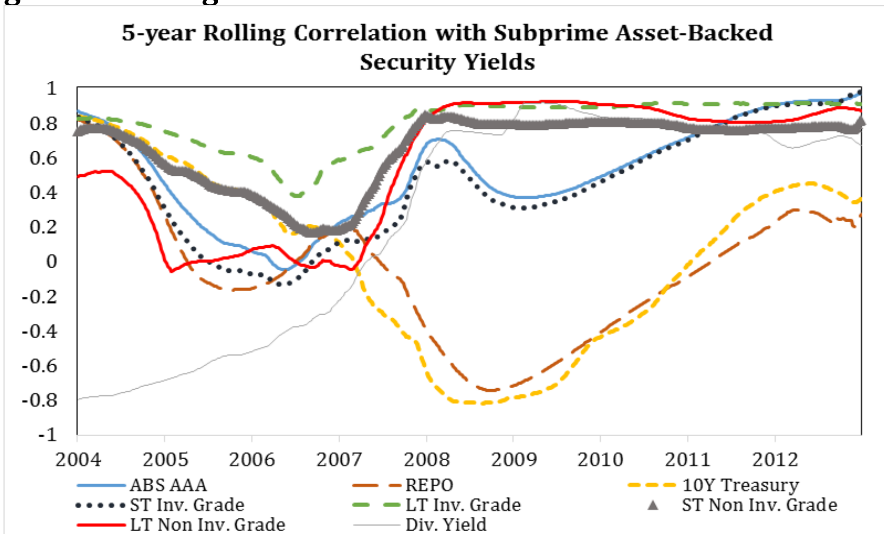


Figure 2: Time Variation of VAR Coefficients in Standard TVP VAR Models

Figure 3: Time Variation of Volatility Coefficients in Standard TVP VAR Models

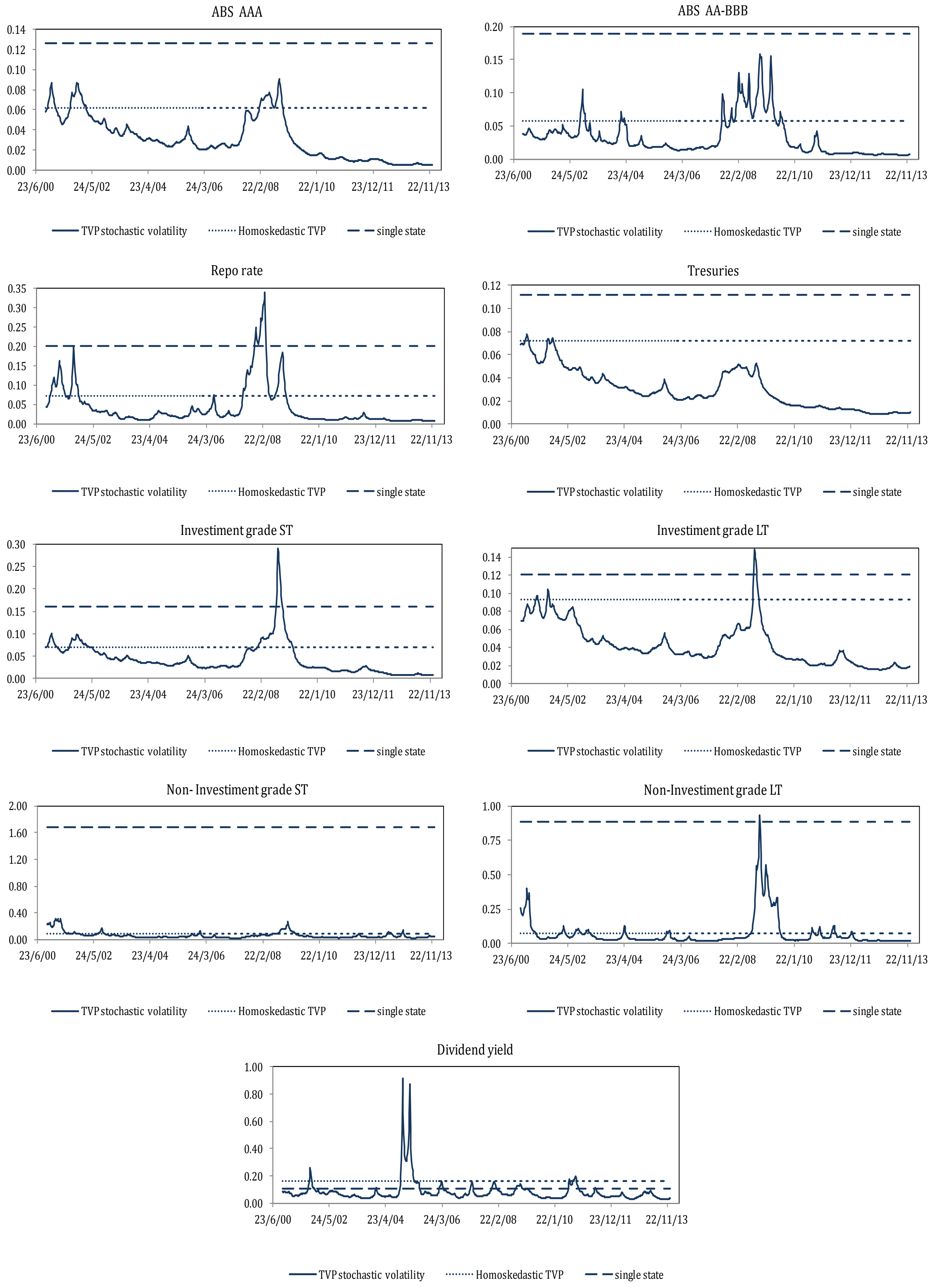


Figure 4: Time Variation of Correlation Coefficients in Standard TVP VAR Models

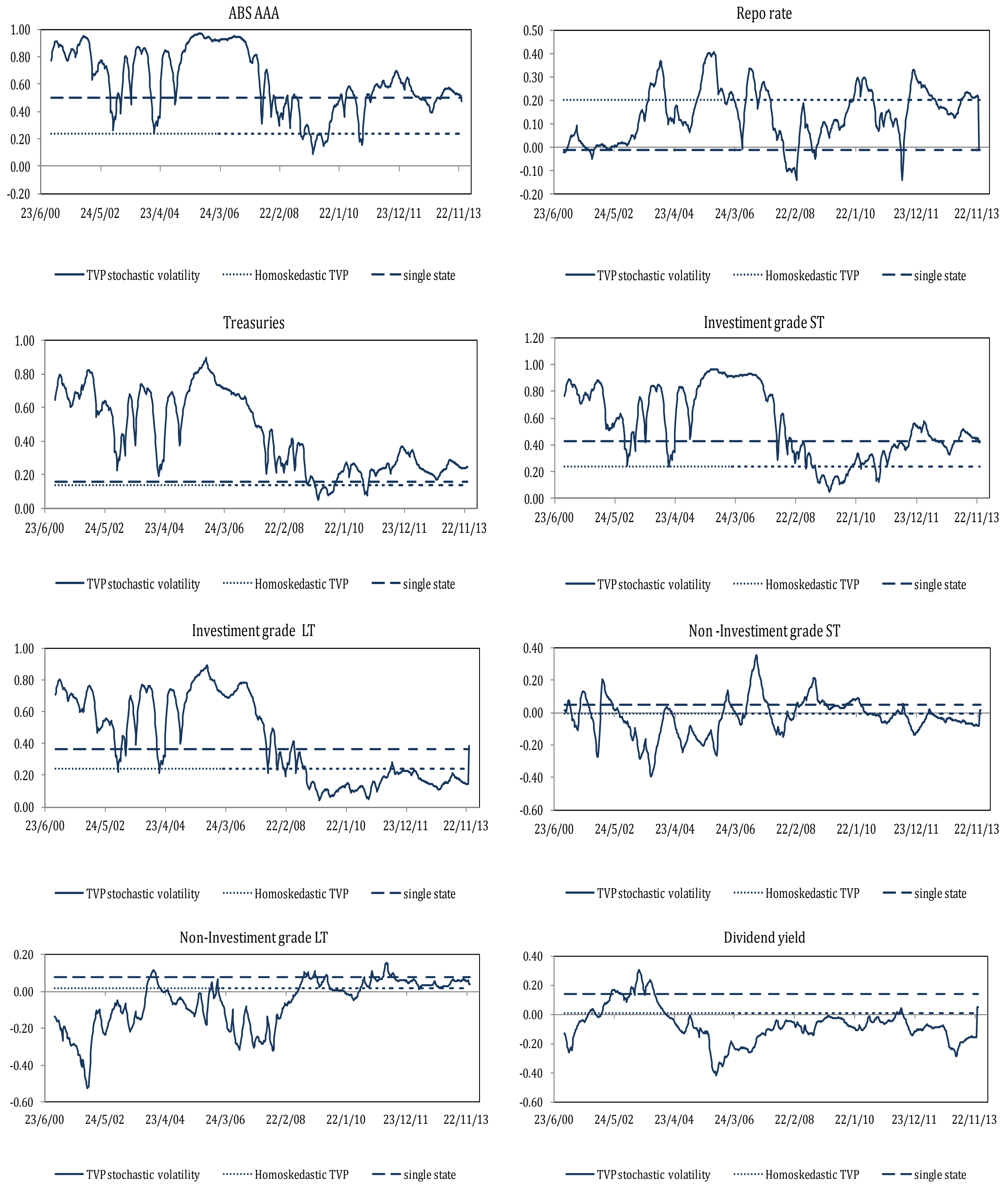


Figure 5: Time Variation in VAR Coefficients in Mixture Innovation TVP VAR Models

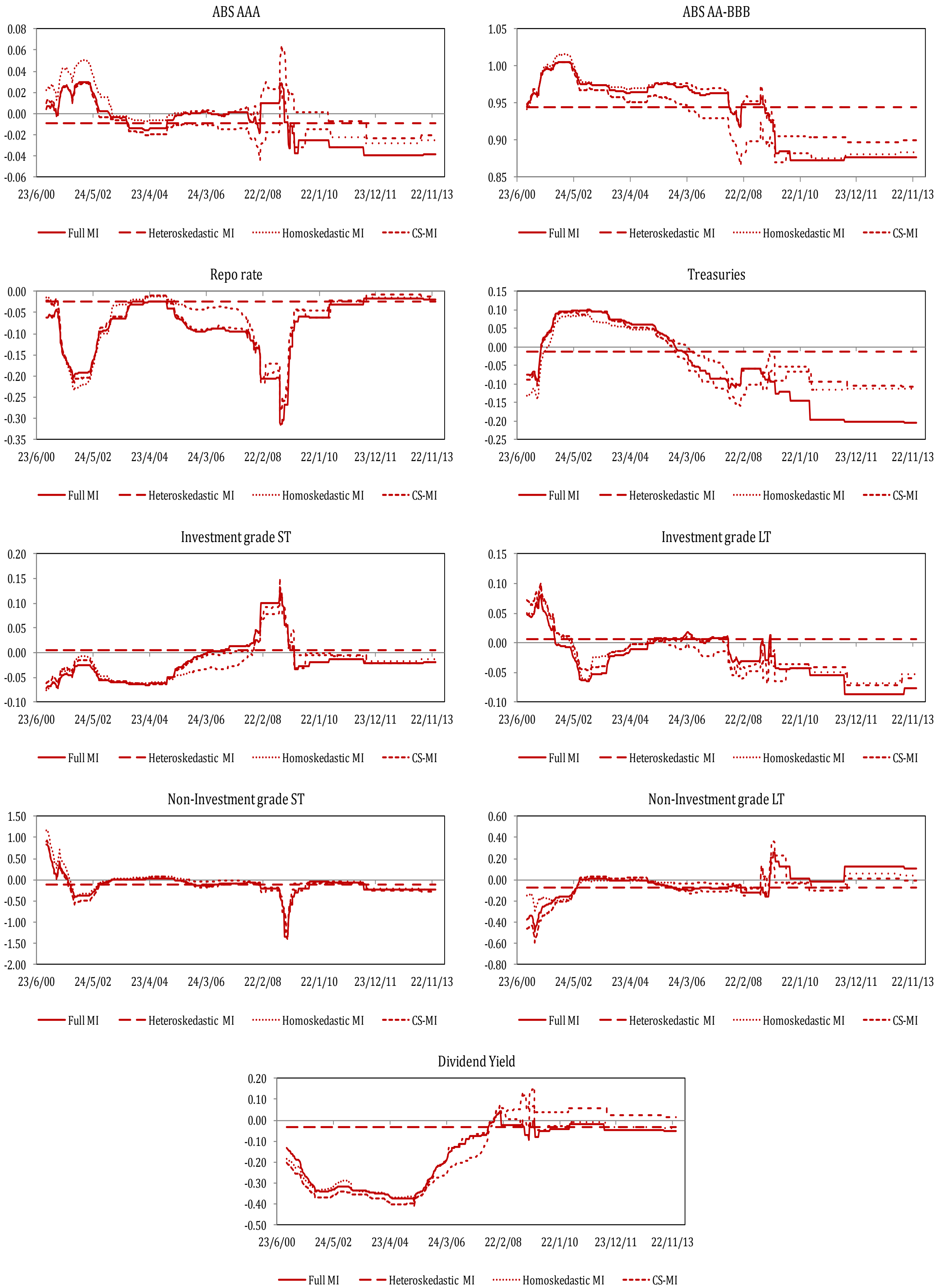
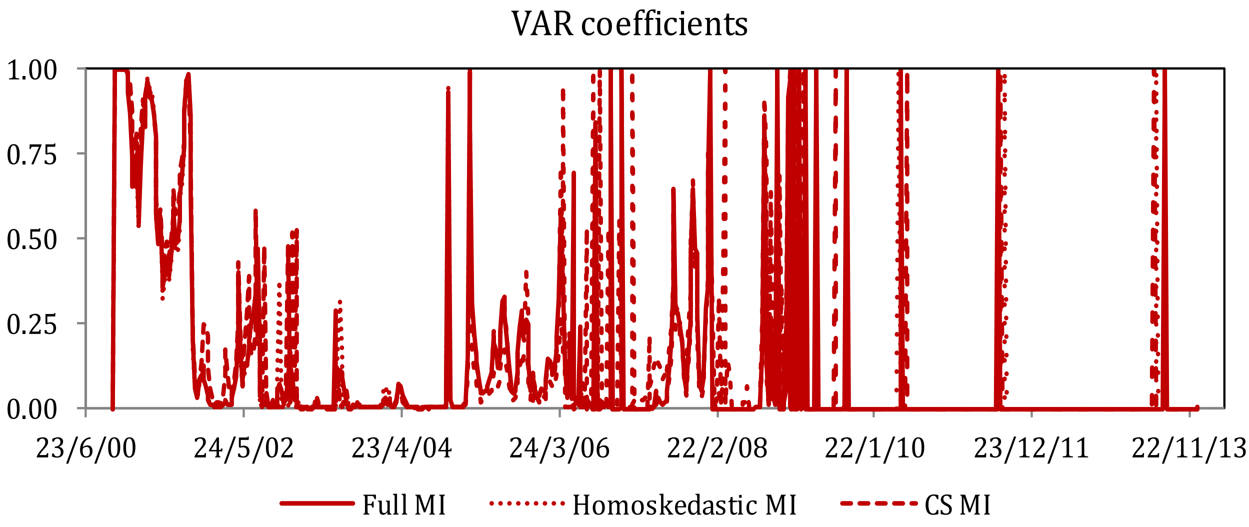
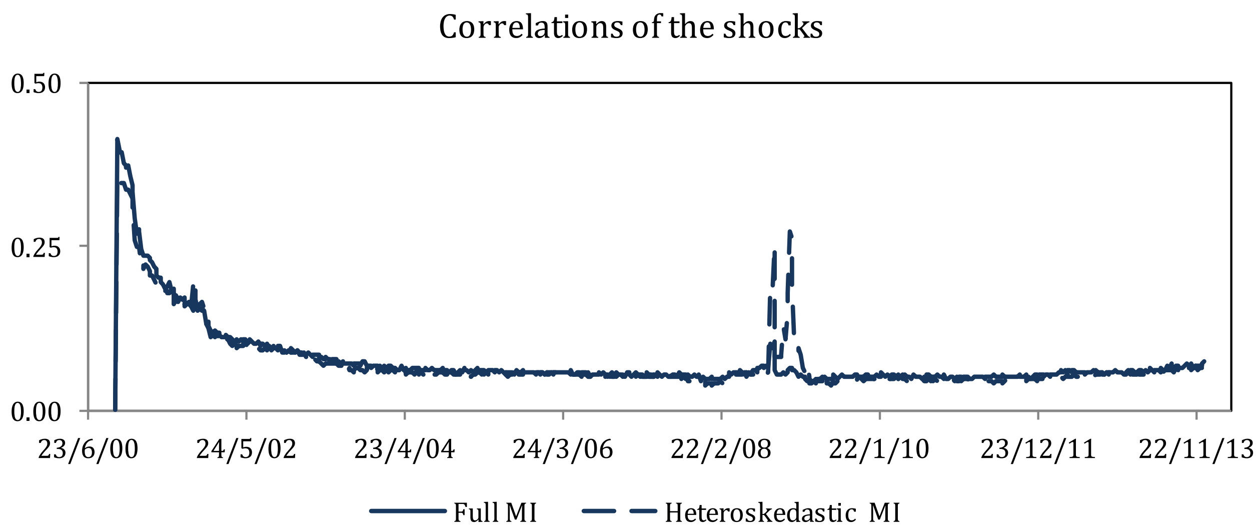


Figure 6: Posterior Estimates of the Probabilities of Breaks from Mixture Innovation TVP VAR Models





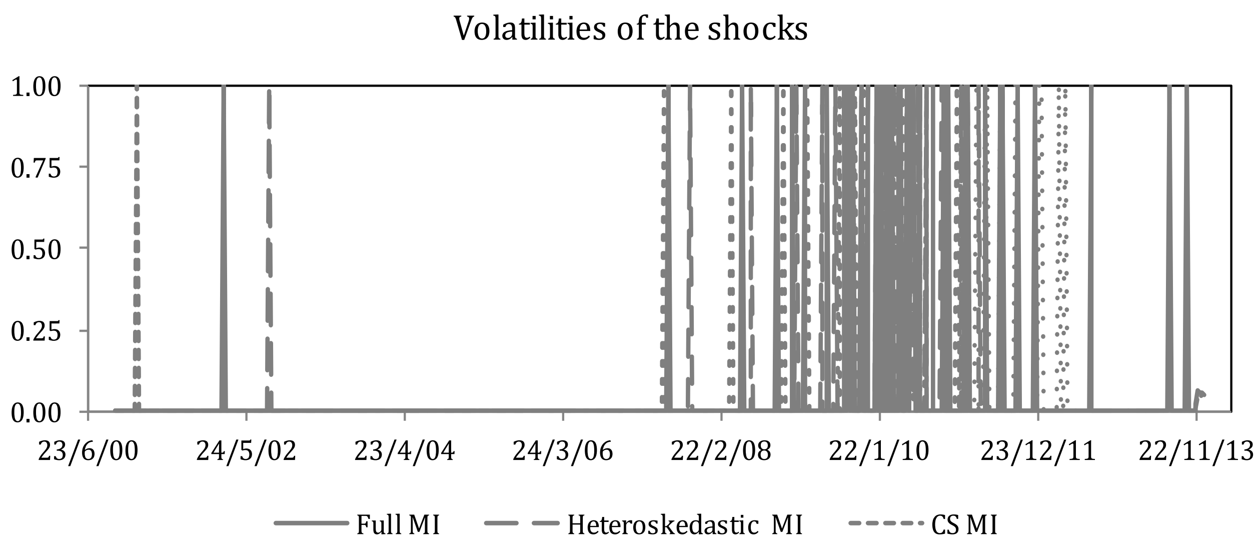


Figure 7: Variation of Volatility Coefficients in Mixture Innovation TVP VAR Models

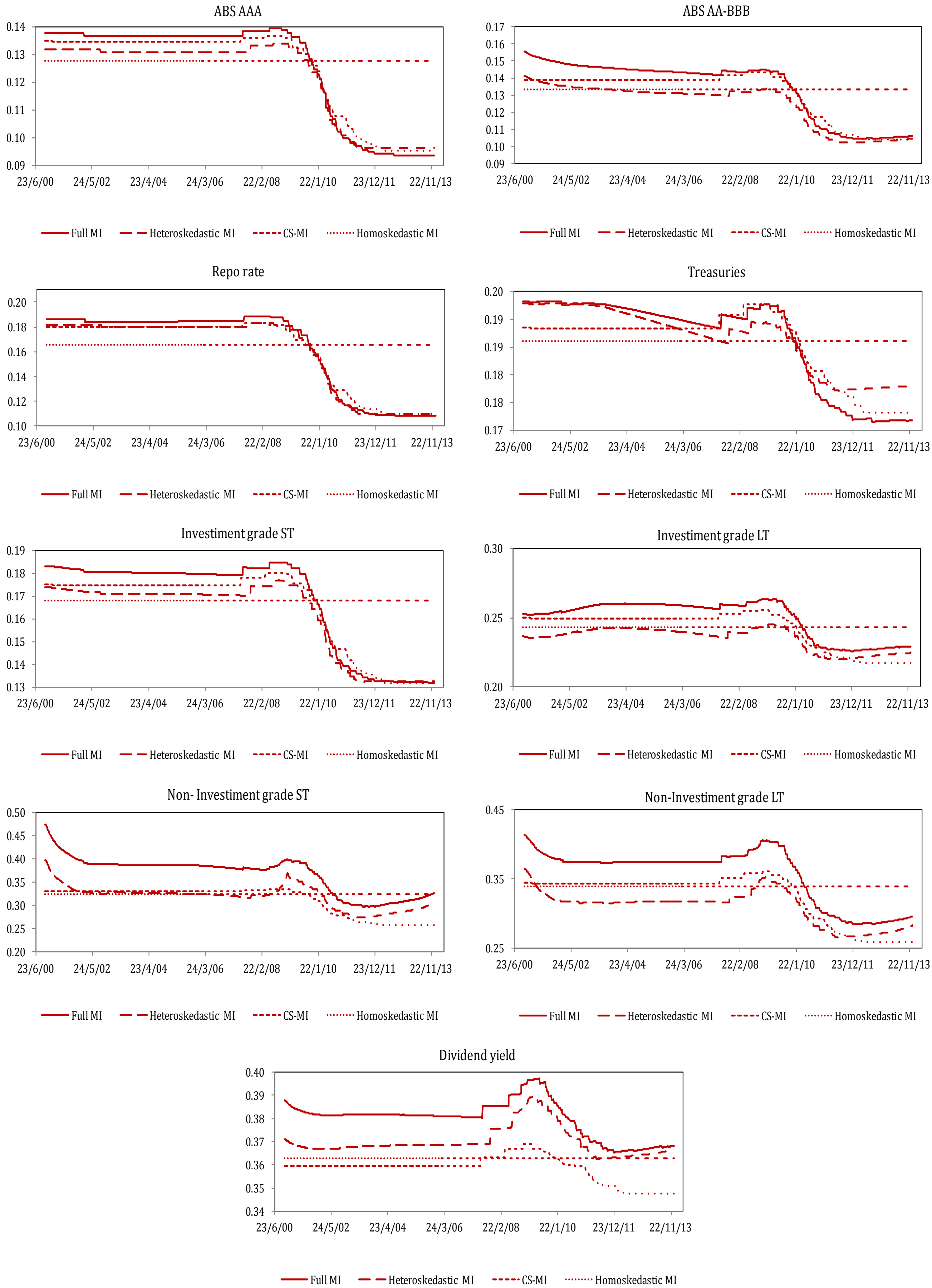


Figure 8: Time Variation of Correlation Coefficients in Mixture Innovation TVP VAR Models

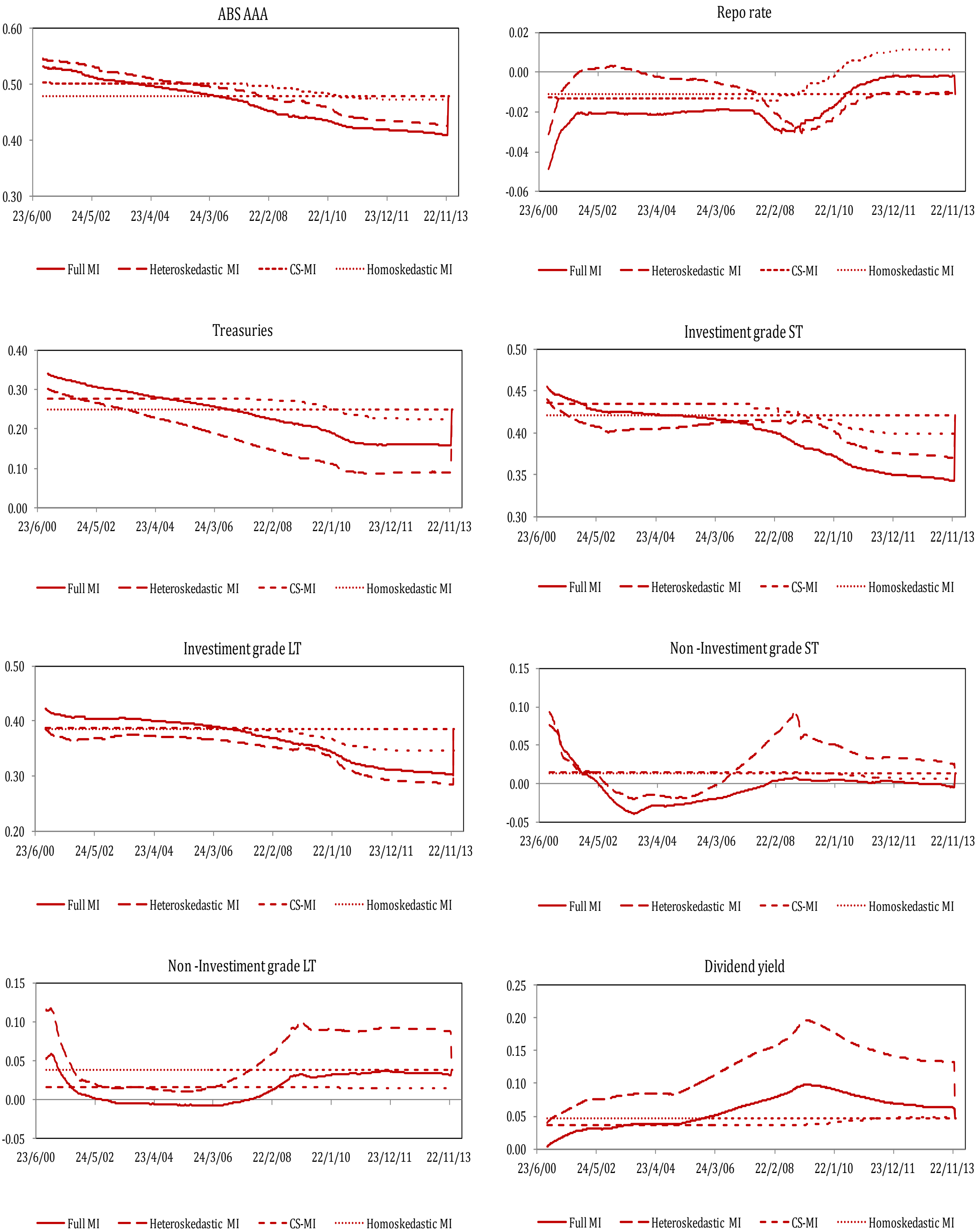


Figure 9: IRF Implied by Homoskedastic TVP VAR –Shock to a Low-Grade ABS Yield

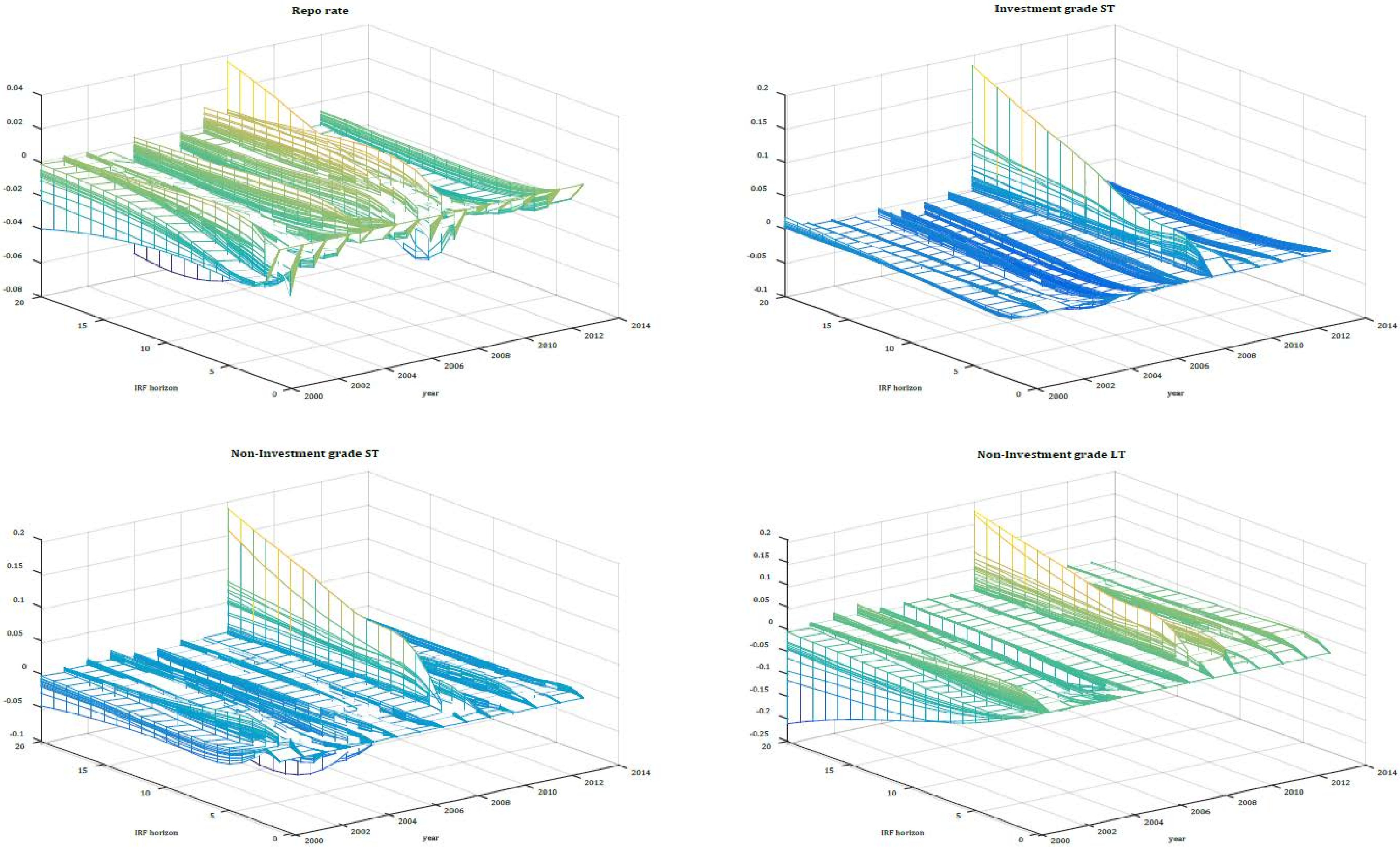


Figure 10: Selected IRFs Implied by Homoskedastic TVP VAR

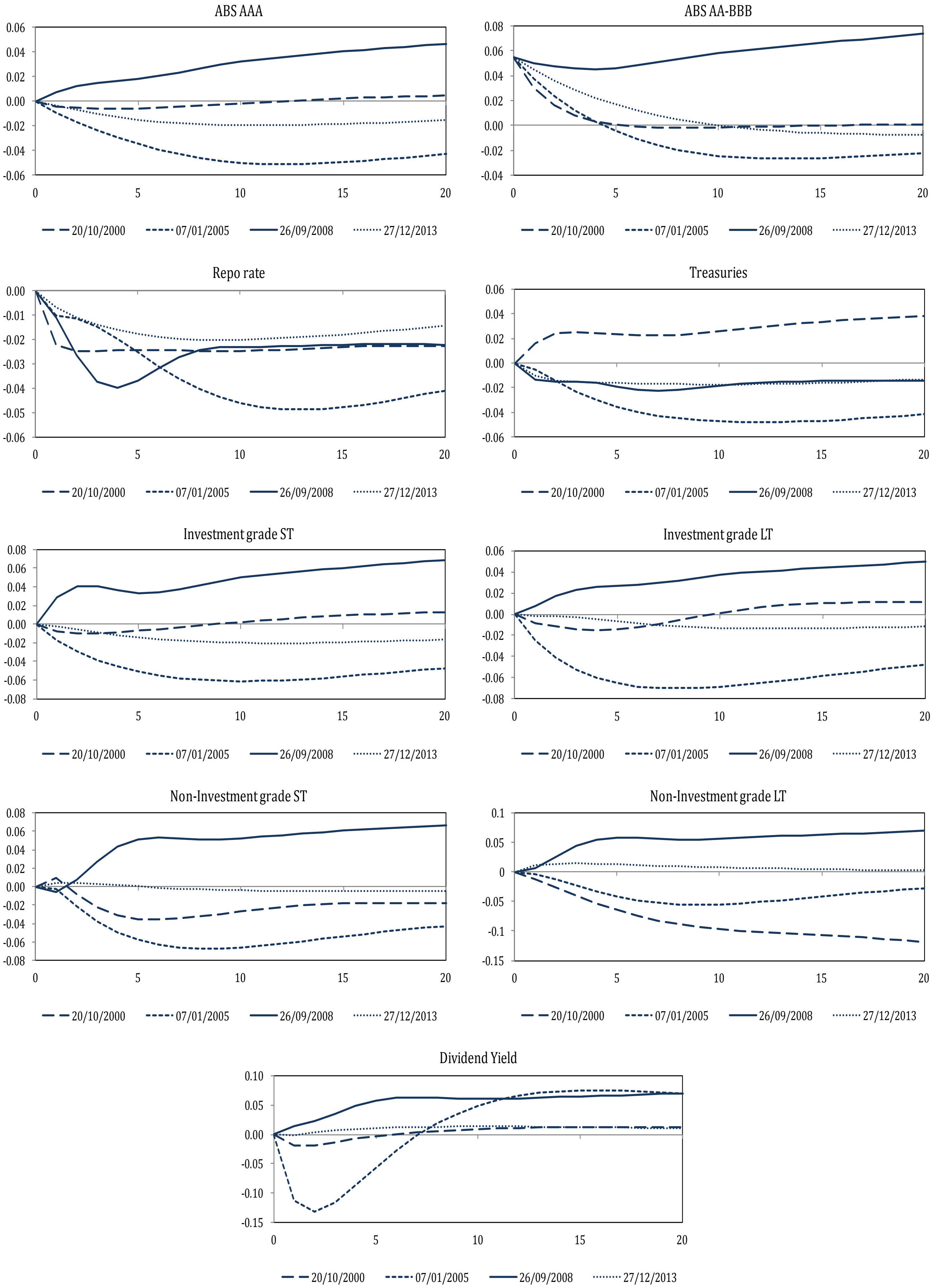


Figure 11: IRF Implied by SV TVP VAR –Shock to a Low-Grade ABS Yield

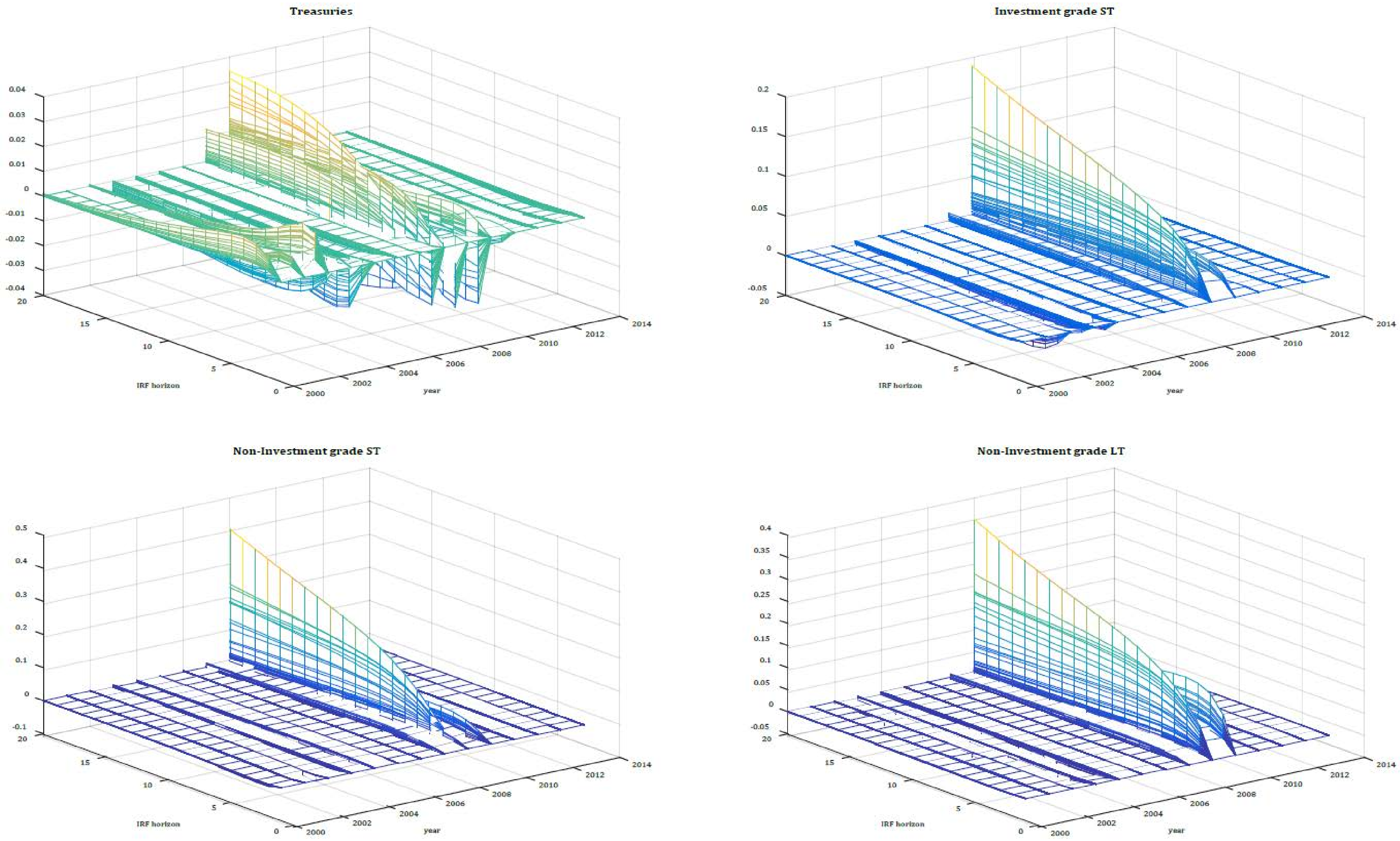


Figure 12: IRF by a Mixture Innovation TVP VAR –Shock to a Low-Grade ABS Yield

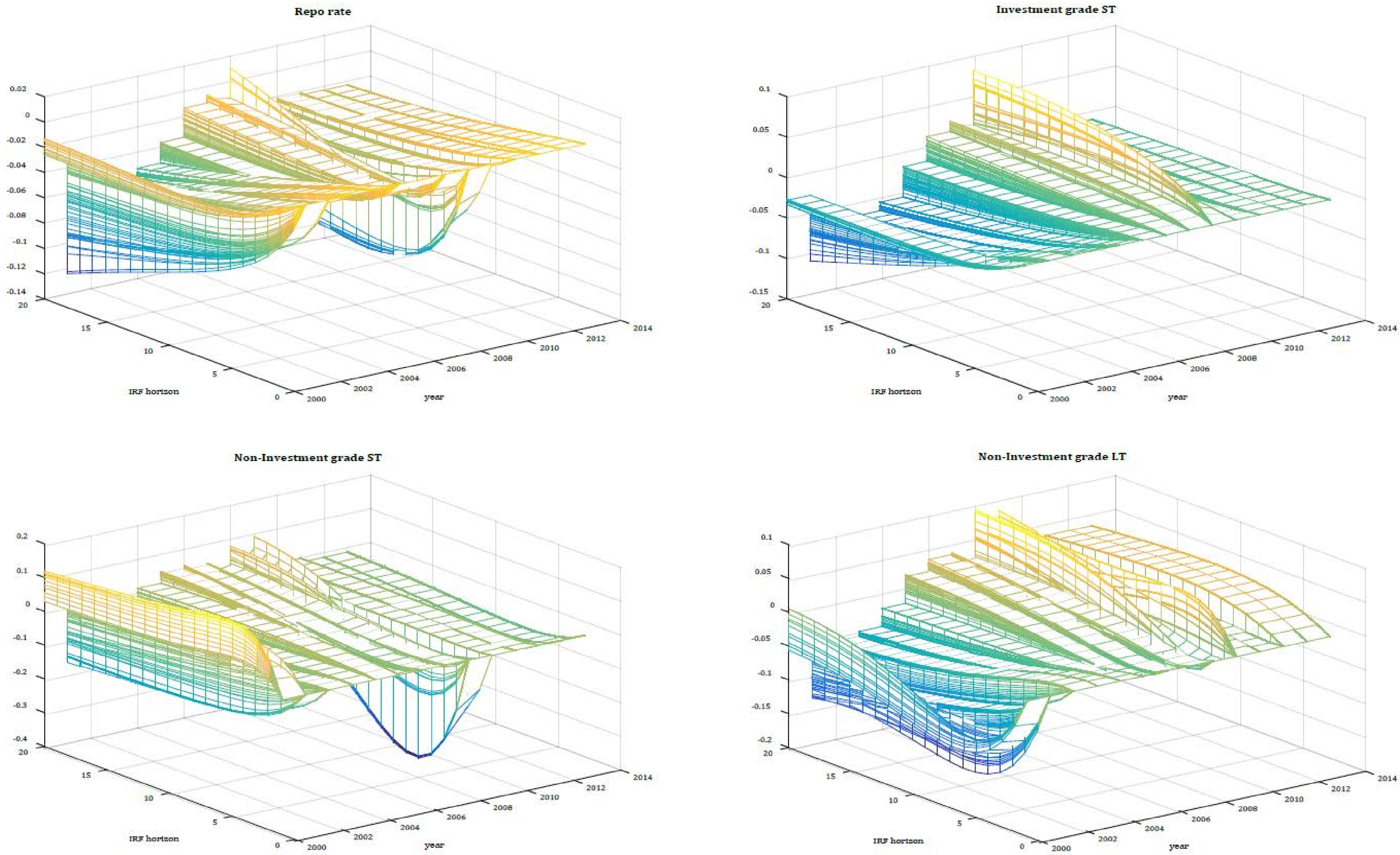


Figure 13: Selected IRFs Implied by a Mixture Innovation TVP VAR Model