Journal of Financial Markets 45 (2019) 83–114

Contents lists available at ScienceDirect

Journal of Financial Markets

journal homepage: www.elsevier.com/locate/finmar

Cross-asset contagion in the financial crisis: A Bayesian time-varying parameter approach☆

Massimo Guidolin [a,*], Erwin Hansen [b], Manuela Pedio [c]

[a] Department of Finance, BAFFI-CAREFIN, and IGIER, Bocconi University, Milan, Italy
[b] Departamento de Administración, Facultad de Economía y Negocios, Universidad de Chile, Address: Diagonal Paraguay 257, Oficina, 1204, Santiago, Chile
[c] Department of Finance, Bocconi University, Milan, Italy

ARTICLE INFO

Article history:
Received 12 December 2017
Received in revised form 5 April 2019
Accepted 8 April 2019
Available online 13 May 2019

JEL classification:
E40
E52

Keywords:
Contagion
Bond yield
Financial crisis
Interdependence
Bayesian estimation

ABSTRACT

The recent U.S. subprime crisis provides us with a perfect framework to study cross-asset contagion mechanisms in the U.S. financial markets. Specifically, 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 to capture the (potentially) time-varying relation among the different financial markets. 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 the correlated information channels.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

The instability in key macroeconomic and financial indicators as well as in their linkages has been extensively studied in the literature (Stock and Watson, 1996; Sensier and Van Dijk, 2004). In particular, when applied to financial time series, such instabilities have inspired a vast empirical literature on contagion, a state of alteration of such linkages that would be characterized by their strengthening in excess to either some notion of normality or of fundamental connections (e.g., Dornbusch, et al., 2000; Ciccarelli and Rebucci, 2003; Pericoli and Sbracia, 2003; Jawadi et al., 2015). However, most researchers have focused on cross-country relationships, in particular, on episodes of international contagion (e.g., King and Wadhwani, 1990; Kodres and Pritsker, 2002; Kaminsky et al., 2003; Dungey et al., 2007). In this study, we perform an empirical investigation of the instability of financial linkages and of contagion effects, but we focus instead on cross-asset,

☆ We thank Tarun Chordia (the editor) and one anonymous referee for constructive comments and encouragement. We are also grateful to Philipp Hartman, Simone Manganelli, Xiang Zhang (a disussant), and seminar/session participants at the European Central Bank, the European Financial Management Associaton 2018 Meetings in Milan, and the Financial Engineering and Banking Society 2018 Meetings in Rome. Erwin Hansen acknowledges financial support from FONDECYT grant # 11150693.

* Corresponding author.
E-mail addresses: massimo.guidolin@unibocconi.it (M. Guidolin), ehansen@fen.uchile.cl (E. Hansen), manuela.pedio@unibocconi.it (M. Pedio).

https://doi.org/10.1016/j.finmar.2019.04.001
1386-4181/© 2019 Elsevier B.V. All rights reserved.

within-country, contagion mechanisms.[1] We exploit the recent U.S. 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 to investors, because identifying contagion and distinguishing from spillover phenomena with sufficient precision would allow them to better adjust their portfolio diversification strategies, and for policy-makers, as the effectiveness of policy interventions is likely to depend on the nature of the shock transmission channels. We propose a novel empirical strategy to test whether the propagation effects from asset-backed securities (ABSs) 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 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. Therefore it is crucial to distinguish episodes of financial contagion from “normal” interdependences among asset classes/markets, whereby the former is reflected by an extreme variation of the model parameters while evidence of constant (or smoothly changing) parameters will capture normal interdependence among assets through time (e.g., Billio and Caporin, 2010). Also, following Ciccarelli and Rebucci (2003), we narrow our definition of contagion to a temporary change in cross-markets linkages in order to distinguish it from permanent changes in the transmission mechanisms, namely structural breaks (also see the discussion in Abbate et al., 2016). We emphasize that our definition of interdependence and contagion differs from more traditional definitions (e.g., Forbes and Rigobon, 2002; Bacchiocchi, 2017) that exploit heteroskedasticity and therefore changes in the contemporaneous cross-correlations of structural shocks, usually in a vector autoregressive (VAR) model. In contrast, we study the dynamics of the structural shocks to differentiate volatility spillovers from pure contagion episodes.

In our study, we put these definitions at work by first tracking the evolution of three blocks of parameters: the VAR coefficients governing the long-term relations between the endogenous variables in our model, the volatility coefficients of the shocks, and the correlations between such exogenous shocks. In particular, shifts in VAR coefficients reflect any non-linear effects of variable co-movements that fit Ciccarelli and Rebucci's definition of contagion, while the volatilities and correlations of the residuals allow us to identify episodes of apparent contagion (or volatility spillovers), 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 et al., 2018). Finally, we study the instability in the transmission mechanisms of the propagation of the shocks by observing how the impulse response functions (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 capture 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). Some researchers have discussed—often reaching conclusions that favor flexible Bayesian methods (e.g., Bianchi, et al., 2017; Tu, 2010) —the possibility that frequentist methods may fail to efficiently handle shifting parameters (unless convenient structure is imposed, as in Eickmeier et al., 2015), a key element to separate out contagion effects from constant-linkage effects in the presence of more variable shocks (e.g., Forbes and Rigobon, 2002). In particular, we propose flexible models that allow for financial instability and cross-asset contagion using Bayesian VAR models with time-varying parameters. We explore several alternative specifications ranging from single-state 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 model with stochastic volatility (SV) where time variation also affects the covariance matrix. Finally, we extend the range of models entertained to mixture innovation (MI) TVP-VAR models similar to those proposed by Gerlach et al. (2000) and Giordani and Kohn (2008). More precisely, this specification eliminates all restrictions imposed by the

[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 studying the particular effects in each of these markets, and of course, their comovement and interdependence, is worthy of attention. For instance, Brenner et al. (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 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 (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 U.S. and eventually the global economy into what we now call the Great Recession. Li et al. (2016) present 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.

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 for the yields 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 MI 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 contagion occurred during periods of financial distress and especially during the Great Financial Crisis (GFC) of 2007–2009. In particular, with refence to the GFC, we find that contagion effects were initially mainly driven by remarkable instability in the endogenous cross-market relationships, while in the last part of the crisis there is also evidence of significant variation in the scale and co-variation of the exogenous 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, flight-to-quality, and correlated information channels (these are explained in Section 5). We find instead only weak evidence of propagation through a well-known risk premium channel, by which in time of distress investors would decrease the amount of risky assets in their portfolios in favor of relatively riskless ones.

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 GFC in the U.S. For instance, Hartmann et al. (2004) use non-parametric extreme value theory methods to study financial market co-movements not only in an international dimension (the G-5 countries), but also across asset classes, with emphasis on stock-bond contagion and fligh-to-quality patterns. They report non-negligible both domestic and international cross-asset market linkages in times of stress. Longstaff (2010) focuses on contagion from the relatively small collateralized debt obligation (CDO) market in the U.S. and analyses the mechanisms through which a negative shock in this market propagates to the other U.S. asset classes (Treasuries, corporate bonds, stocks, and the VIX). However, Longstaff inspects potential changes in the linkages across these markets by separately and exogenously analyzing the pre-crisis, the subprime crisis, and the global-crisis period. In contrast, 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) extend the analysis in Longstaff (2010) using a Markov switching (MS) time-varying probabilities VAR model. Their results surprisingly suggest that cross-asset contagion was weak during the recent subprime crisis, which was mainly driven by stable interdependences across the U.S. markets. Finally, Guo et al. (2011) study the contagion effects using a Markov switching VAR (MSVAR) 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, researchers have also studied contagion effects during the recent European debt crisis. Bacchiocchi (2017) proposes a 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. Kohonen (2014) studies the transmission of government default risk in the eurozone using a SVAR model that allows for structural breaks in the intercepts (which are interpreted as exogenous changes in idiosyncratic country risk), in the autoregressive coefficients (which would capture changes in the interdependence patterns among countries and referred to as spillovers effects), and in the contemporaneous cross-correlations, which might capture contagion in times of crisis. In contrast to these two studies, in this paper we identify contagion by looking at the sudden changes of the coefficient matrix of a VAR.

A second literature strand that is related to our work studies 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 (e.g., Ciccarelli and Rebucci, 2007; Arakelian and Dellaportas, 2012; Bai et al., 2012; Kaabia et al., 2013; Casarin et al., 2015).

[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; Hong et al., 2006).

[4] As explained in Section 4, 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.

[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).

The rest of the paper is organized as follows. In Section 2, we describe the econometric models and the Bayesian estimation methods. In Section 3, we present the data and descriptive statistics. In Section 4, we report our main empirical results. In Section 5, we evaluate the alternative transmission channels of contagion across asset classes. In Section 6, we report a set of robustness checks, especially concerning the role of priors. We conclude in Section 7.

2. Methodology

In this section, we describe our general econometric framework and discuss estimation issues and the computation of IRFs for alternative models. Similarly to Primiceri (2005) and Koop and Korobilis (2010), we use time-varying-parameter VAR models 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

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

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

where $\mathbf{y}_t$ and $\mathbf{u}_t$ are $n \times 1$ vectors containing, respectively, the dependent variables and the error terms in the measurement equation (1) that captures how (say) asset yields depend on a 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 a $n \times p_0$ matrix of predetermined explanatory variables $\mathbf{W}_t$ into the dependent variables; $\boldsymbol{\beta}_t$ is a $m \times 1$ vector of time-varying coefficients that multiplies the $n \times m$ matrix of additional explanatory variables $\mathbf{Z}_t$. We assume that 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. Crucially, $\boldsymbol{\Omega}_t$ 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, for example, as in Mumtaz and Surico (2009) analysis of international shock transmission in large-scale VARs. Despite the multivariate normality assumption for the shocks in (1)-(2), TVP models are flexible enough to accommodate highly non-normal shapes of the (unconditional) distribution of the variables of interest.

Broadly speaking, in a Bayesian framework, our goal is to obtain from a given sample $\mathbf{y}^T \equiv [\mathbf{y}_1' \ \mathbf{y}_2' \ \dots \ \mathbf{y}_T']'$ the posterior densities of the unknown coefficients, $\boldsymbol{\delta}$, $\boldsymbol{\beta}_t$, $\boldsymbol{\Pi}_t$, $\mathbf{Q}_t$, and $\boldsymbol{\Omega}_t$, to be able to draw inferences on them and estimate meaningful quantities of interest, such as IRFs. In a Bayesian set up, this is performed by first setting up adequate (economically sensible) priors on the matrices $\boldsymbol{\delta}$, $\boldsymbol{\beta}_t$, $\boldsymbol{\Pi}_t$, $\mathbf{Q}_t$, and $\boldsymbol{\Omega}_t$, and then combining these with the evidence contained in the data, as summarized by the likelihood function. Such a task is performed by applying MCMC algorithms explained in the following. However, before proceeding in this way and also to gain some additional economic insights, 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:

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

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

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

[6] In this case, one obtains: $\mathbf{y}_t = \mathbf{c}_t + \mathbf{B}_{1,\,t}\mathbf{y}_{t-1} + \mathbf{B}_{2,\,t}\mathbf{y}_{t-2} + \dots + \mathbf{B}_{q,\,t}\mathbf{y}_{t-q} + \mathbf{u}_t$ 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 by the Gibbs procedure that imply a violation of stationarity are dealt with by rejecting the entire path, to be re-sampled.

Economically, a TVP-VAR model is one in which, because of homoskedasticity, the scale of the shocks is constant over time (and controlled by the fixed matrix $\mathbf{\Omega}$), while each period is characterized by structural instability in the coefficients mapping past values of the yields into future ones. This is caused by the fact that although $E[\boldsymbol{\beta}_t] = \boldsymbol{\beta}_{t-1}$, $\Pr(\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1}) = 0$ except for a null measure event. 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 let the data judge the empirical merits of this model.

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

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

where $\varphi(\cdot)$ represents the Gaussian PDF function. Additionally, we also select the following priors:

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

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 is typical in many papers in the empirical finance and macroeconomics literature (e.g., Koop and Korobilis, 2010). Under the assumption of independent priors, $p(\boldsymbol{\beta}^T, \mathbf{\Omega}, \mathbf{Q}) = p(\boldsymbol{\beta}^T)p(\mathbf{\Omega})p(\mathbf{Q})$, we apply the Gibbs sampler to sequentially draw from the conditional posteriors $p(\mathbf{Q}^{-1}\middle|\mathbf{y}^T, \boldsymbol{\beta}^T)$, $p(\mathbf{\Omega}^{-1}\middle|\mathbf{y}^T, \boldsymbol{\beta}^T)$, and $p(\boldsymbol{\beta}^T\middle|\mathbf{y}^T, \mathbf{\Omega}, \mathbf{Q})$ following the iterative algorithm:

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

At this point, for given posteriors for $\mathbf{\Omega}$ 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 i.i.d. Normal $(\mathbf{0}, \mathbf{\Omega}_t)$, in which the covariance matrix of error terms $\mathbf{\Omega}_t$ is time-varying. We also 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 modelled as follows:

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

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

where $u_{i,t}$ and $\eta_{i,t+s}$ are independent of 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 in (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 refer to equation (9) as the log-volatility process.

[7] We refer to Harrison and West (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 (SV univariate 1).

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

$$\mathbf{y}_t = \mathbf{c}_t + \mathbf{B}_{1,\,t}\mathbf{y}_{t-1} + \mathbf{B}_{2,\,t}\mathbf{y}_{t-2} + \ldots + \mathbf{B}_{q,\,t}\mathbf{y}_{t-q} + u_t \quad \text{with} \quad u_t \sim N(\mathbf{0},\ \boldsymbol{\Omega}_t) \tag{10}$$

$$\mathbf{A}_t\boldsymbol{\Omega}_t\mathbf{A}_t' = \boldsymbol{\Sigma}_t\boldsymbol{\Sigma}_t', \tag{11}$$

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

We can re-write the model (10)-(11) in compact form as:

$$\mathbf{y}_t = \mathbf{Z}_t'\boldsymbol{\beta}_t + \mathbf{A}_t^{-1}\boldsymbol{\Sigma}_t\boldsymbol{\varepsilon}_t \quad \text{with } \boldsymbol{\varepsilon}_t \sim N(\mathbf{0},\ \mathbf{I}_n) \tag{12}$$

$$\mathbf{Z}_t' \equiv \mathbf{I}_n \otimes \left[1\ \mathbf{y}_{t-1} \ldots \mathbf{y}_{t-q}\right],\ \boldsymbol{\beta}_t \equiv \left[\mathbf{c}_t\ \mathbf{B}_{1,\,t} \ldots \mathbf{B}_{q,\,t}\right]. \tag{13}$$

Following Primiceri (2005), we stack the unrestricted elements of the matrices $\mathbf{A}_t$ and $\boldsymbol{\Omega}_t$ in the vectors $\mathbf{a}_t$ and $\boldsymbol{\sigma}_t^2$, respectively, and assume that the corresponding coefficients follow the processes:

$$\begin{aligned} &\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\nu}_t \\ &\mathbf{a}_t = \mathbf{a}_{t-1} + \boldsymbol{\zeta}_t \\ &\ln\left(\boldsymbol{\sigma}_t^2\right) = \ln\left(\boldsymbol{\sigma}_{t-1}^2\right) + \boldsymbol{\eta}_t, \end{aligned} \tag{14}$$

where $\ln(\boldsymbol{\sigma}_t^2)$ indicates the element-wise natural log of the elements of $\boldsymbol{\sigma}_t^2$. 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 $\mathbf{Q}$, $\mathbf{S}$, and $\mathbf{W}$ positive definite matrices.[9] From an economic perspective, the model in (10)-(11) forces all parameters, in this case also the variances and (indirectly, through the $\boldsymbol{\Omega}_t = \mathbf{A}_t'\boldsymbol{\Sigma}_t\boldsymbol{\Sigma}_t'\mathbf{A}_t$ transformation) covariances, to change in correspondence to all points in the sample. Of course, by making the covariance matrix of all shocks $(\boldsymbol{\nu}_t,\ \boldsymbol{\zeta}_t,\ \boldsymbol{\eta}_t)'$ sufficiently small, it remains possible to capture situations in which parameters evolve smoothly with small period-to-period changes, apart from occasionally large 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

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

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 $\widehat{\mathbf{A}}_{ii,\ OLS}$ denote the ith block of $\mathbf{S}$ and the correspondent block of $\widehat{\mathbf{A}}_{OLS}$, respectively. $\widehat{\mathbf{A}}_{OLS}$ can be computed from the relationship $\mathbf{A}Cov(\widehat{\mathbf{u}}_t^{OLS})\mathbf{A}' = diag(\widehat{\boldsymbol{\sigma}}_{OLS}^2)$. Based on the selected priors and conditionally on the data observations, the Gibbs

[9] $\mathbf{S}$ 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 replaced by non-zero blocks; however, we opt for this specification in order to reduce the number of parameters in the model and allow for a structural interpretation of the innovations.

sampler performs posterior simulations sequentially drawing the VAR coefficients $\boldsymbol{\beta}^T$, the simultaneous relations $\mathbf{A}^T \equiv [\mathbf{A}_0\ \mathbf{A}_1\ \mathbf{A}_2 \ldots \mathbf{A}_T]'$, the volatilities $\boldsymbol{\Sigma}^T \equiv [\boldsymbol{\Sigma}_0 \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2 \ldots \boldsymbol{\Sigma}_T]'$, 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. $p(\boldsymbol{\beta}^T | \mathbf{y}^T, \mathbf{A}^T, \boldsymbol{\Sigma}^T, \mathbf{V})$.

2.4. TVP VAR with mixture innovations

The pioneering model presented by Primiceri (2005) 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 also consider 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 is a model of infrequent, possibly rare, parameter shifts that can indeed trigger periods of altered co-movements that we normally denote 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 control 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 ($\boldsymbol{\Sigma}_t$), respectively. This occurs according to the following dynamics:

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

where $K_{it} \in \{0, 1\}$ for $t = 1, 2, 3$. Crucially, the breaks are not restricted to occur 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 the K_{1t}, K_{2t}, and K_{3t} variables 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 scales, the covariances, 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 $\ln(\boldsymbol{\sigma}_t^2) = \ln(\boldsymbol{\sigma}_{t-1}^2)$, which is the case when 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 $\ln(\boldsymbol{\sigma}_t^2) = \ln(\boldsymbol{\sigma}_{t-1}^2)$: no contagion or volatility spillovers, just normal interdependence.
2) $\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\nu}_t$, $\mathbf{a}_t = \mathbf{a}_{t-1}$, and $\ln(\boldsymbol{\sigma}_t^2) = \ln(\boldsymbol{\sigma}_{t-1}^2)$: 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 $\ln(\boldsymbol{\sigma}_t^2) = \ln(\boldsymbol{\sigma}_{t-1}^2)$: 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 $\ln(\boldsymbol{\sigma}_t^2) = \ln(\boldsymbol{\sigma}_{t-1}^2) + \boldsymbol{\eta}_t$: 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 $\ln(\boldsymbol{\sigma}_t^2) = \ln(\boldsymbol{\sigma}_{t-1}^2)$: both contagion and volatility spillover at work, from altered interdependence.[10]

Economically, each of these cases is important. As unlikely as it may appear in our application, case 1 represents the null hypothesis of most analyses that show no concern for instability in financial relationships. Case 2 perfectly fits the definition of contagion we have adopted following the literature, but appears to be empirically restrictive because it is written jointly with an assumption of homoskedasticity. Cases 3 and 4 are, at least in an economic sense, our alternative hypotheses because under these processes, the data may show behavior that is often addressed as if it represented an instance of contagion, while we can just detect altered volatility and/or covariance spillovers. However, case 5 represents the really valuable and delicate one: when the data contain evidence of both volatility and covariance spillovers and of genuine contagion, and only an econometric model provides a chance to tell the two apart.

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 evolution of

[10] The case of $\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1}$, $\mathbf{a}_t = \mathbf{a}_{t-1} + \boldsymbol{\zeta}_t$, and $\ln(\boldsymbol{\sigma}_t^2) = \ln(\boldsymbol{\sigma}_{t-1}^2) + \boldsymbol{\eta}_t$; $\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\nu}_t$, $\mathbf{a}_t = \mathbf{a}_{t-1}$, and $\ln(\boldsymbol{\sigma}_t^2) = \ln(\boldsymbol{\sigma}_{t-1}^2) + \boldsymbol{\eta}_t$ are of course possible. They occur when a crisis causes both altered interdependence and/or contagion, as well as higher volatility than in normal times.

the states. In Section 4, we estimate all the variants of the baseline model (1)–(5) also to try and economize over the complexity implied.

One important restriction of the general mixture model in Primiceri (2005) and Koop et al. (2009) is represented by what we call the Cogley-Sargent variant (2005, CG), 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 call 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 equation (4) above, elevated to the rank of a separate model because the restriction is imposed ex-ante, as if an infinitely precise prior were imposed.

Model estimation is performed including a few additional steps and cautions in the Gibbs sampler that are becoming standard in the literature (e.g., Gerlach et al., 2000).[11] In particular, we need to define the prior distribution of the vector $\mathbf{K}_t$ 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 $Beta\left(\underline{\beta}_{1i}, \underline{\beta}_{2i}\right)$ is used for the change-point probabilities p_1, p_2, and p_3. As a result, the conditional posterior for p_i, which is 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, which assumes breakpoints in $\boldsymbol{\beta}_t$, $\mathbf{a}_t$, and $\ln(\boldsymbol{\sigma}_t^2)$ to be independent of one another, we can individually generate K_{1t}, K_{2t}, and 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 in Subsections 2.2-2.3. As for the additional mixture innovation parameters, we opt for an informative "few-breaks" Beta prior $B\left(\underline{\beta}_{1j}, \underline{\beta}_{2j}\right)$ with $\underline{\beta}_{1j} = 0.1$ and $\underline{\beta}_{2j} = 9.37$, for all $j = 1, 2, 3$. According to the properties of the Beta distribution, this prior implies a break probability of 1.056% per *period* only. This guarantees that any evidence of frequent breaks comes not from our willingness to detect them, but from the data, through the likelihood function.[12] In Section 6, we performs a number of robustness checks concerning the choice of the prior probability of recording breaks and its effects on our key empirical findings.

2.5. Markov switching VAR

Finally, because this model has often appeared in the literature on cross-asset contagion (e.g., Connolly et al., 2007; Guo et al., 2011; Flavin and Sheenan, 2015; Guidolin and Pedio, 2017), 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:

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

$S_{1t} = 0, 1$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

[11] To use the Gibbs sampler within the new TVP-VAR specification, we only need to make a few minor adjustments. First, 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 $\mathbf{W}$ do 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 $i = 1, 2, 3$).

[12] Note that a break probability of 1.056% per period implies a probability of recording at least one break over T observations that equals $1 - (1 - 0.01056)^T$; for $T = 656$, such a probability is 0.999. Such a value of 1.056% per month has been obtained by multiplying by 10 the per-period probability that over 656 observations, one would end up with a probability of at least one break of 0.5, that equals $0.5 = 1 - (1 - \pi)^{656}$; this can be solved to yield $\pi = 0.001056$. A $\times 10$ value $\pi = 0.01056$ strikes a compromise between a relatively negligible probability of a break in each single period and almost surely recording at least one break over our overall sample. We thank an anonymous referee for stimulating us to work on plausible priors not excessively averse to the chances of witnessing to breaks.

variable that drives regime shifts in variance and covariances of the shocks.[13] 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 relies on the idea 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 (because labeling 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 to occur over time without having a recurring nature because $\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\nu}_t$, $\mathbf{a}_t = \mathbf{a}_{t-1} + \boldsymbol{\zeta}_t$ and $\ln(\boldsymbol{\sigma}_t^2) = \ln(\boldsymbol{\sigma}_{t-1}^2) + \boldsymbol{\eta}_t$ 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 $\ln(\boldsymbol{\sigma}_0^2)$ to $\mathbf{a}_1$ and $\ln(\boldsymbol{\sigma}_1^2)$. Although this may appear naïve when compared to mixture TVP-VARs, MSVARs capture in a very intuitive way the idea that coefficients will be unstable but also characterized by rather infrequent shifts, while making it possible to adopt a much simpler estimation approach versus models characterized by shifts governed by $\mathbf{K}_t \equiv [K_{1t} K_{2t} K_{3t}]'$.

Estimation is performed by setting the initial values of the parameters to equal the MLE estimates obtained using the standard expectation-maximization algorithm and then sequentially going through the steps of the Gibbs sampler. The initial state is set to match the vector of ergodic state probabilities of the process. Kim and Nelson (1999) and Frühwirth-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 applied to the 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:

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

The restrictions implicit in $\boldsymbol{\Upsilon}_t = \mathbf{A}_t^{-1} \boldsymbol{\Sigma}_t$ provide the identifying assumptions to isolate the structural shocks so that 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, $\boldsymbol{\Sigma}_t$)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 us 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 $\boldsymbol{\Omega}_t$.[14]

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 $s(s = 1, 2, \ldots, n)$ is defined as:

$$IRF_{t,H}(\mathbf{S}^t) = E\left[y_{i,t+H} | \mathbf{y}^t, \mathbf{S}^t, \boldsymbol{\varepsilon}_t + \nabla_{\boldsymbol{\varepsilon}}\right] - E\left[y_{i,t+H} | \mathbf{y}^t, \mathbf{S}^t, \boldsymbol{\varepsilon}_t\right] \quad i = 1, 2, \ldots, n, \tag{20}$$

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 $\nabla_{\boldsymbol{\varepsilon}}$ is the $n \times 1$ vector of structural shocks given to the system in the experiment (possibly a vector with zeros everywhere but in the sth position). Generally speaking, a generalized impulse response function (GIRF) estimates the differences 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 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

[13] The first-order nature and homogeneity property guarantee that each Markov chain is characterized by constant transition probabilities. Effectively, the two states compound to four regimes with transition parameters: $p_{00}^1 \equiv \Pr(S_{1t} = 0 | S_{1t-1} = 0)$, $p_{11}^1 \equiv \Pr(S_{1t} = 1 | S_{1t-1} = 1)$, $p_{00}^2 \equiv \Pr(S_{2t} = 0 | S_{2t-1} = 0)$, and $p_{11}^2 \equiv \Pr(S_{2t} = 1 | S_{21t-1} = 1)$. 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.

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

available, we compute each impulse response function depending on the initial regime, as well as on the future regimes $\mathbf{S}^H \equiv [S_{1t+1}S_{2t+1}\ldots S_{1t+H}S_{2t+H}]'$ as follows:

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

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 us to know the initial regime:

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

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

3. Data and summary statistics

Our data consist of time series of ex-ante yields of several asset classes traded in the U.S. financial market. In particular, we collect information on 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 5103 observations over the nine series.

To capture the onset of the U.S. financial crisis and an initial shock derived from the asset-backed securities, 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, which are backed by mortgages. Before the subprime crisis, ABS came to represent popular investments, as they allowed banks and other institutions subject to capital requirements to reduce the size of their balance sheets and free up capital. In 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. Today, 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 U.S. 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 us 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 U.S. 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.[15]

Table 1 (Panel A) presents the main statistical features of our data (while plots are available upon request). Sample means and standard deviations are relatively unsurprising: ABS have 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, these yields are the

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

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	Kurtosis	Skewness	Jarque-Bera
ABS AAA	3.65	−1.04	0.20	38.00
ABS AA-BBB	6.64	4.91	2.01	1207.08
Repo rate	2.12	−0.95	0.70	94.85
Treasuries	3.91	−0.44	−0.19	10.50
Investment grade ST	3.60	−0.88	0.34	38.06
Investment grade LT	6.43	−0.46	0.19	10.78
Non-Investment grade ST	20.40	3.59	1.75	754.49
Non-Investment grade LT	12.78	7.46	2.59	2472.56
Dividend Yield	1.92	1.74	0.67	144.07

	ABS AAA	ABS AA-BBB	Repo rate	Treasuries	Inv. grade ST	Inv. grade LT	Non-Inv. grade ST	Non-Inv. grade LT	Dividend Yield
Panel B – Sample correlations									
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

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 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 with 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 the pricing of the corresponding bonds during long portions of our sample leading up to the financial crisis.

In Fig. 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 Fig. 1 that the subprime shock in the low-grade ABS market did spill over to all asset classes in the U.S. financial market by affecting the joint dynamics of their yields. This preliminary, unconditional, evidence motivates our empirical investigation of interdependence and contagion among asset classes that follows.

4. 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

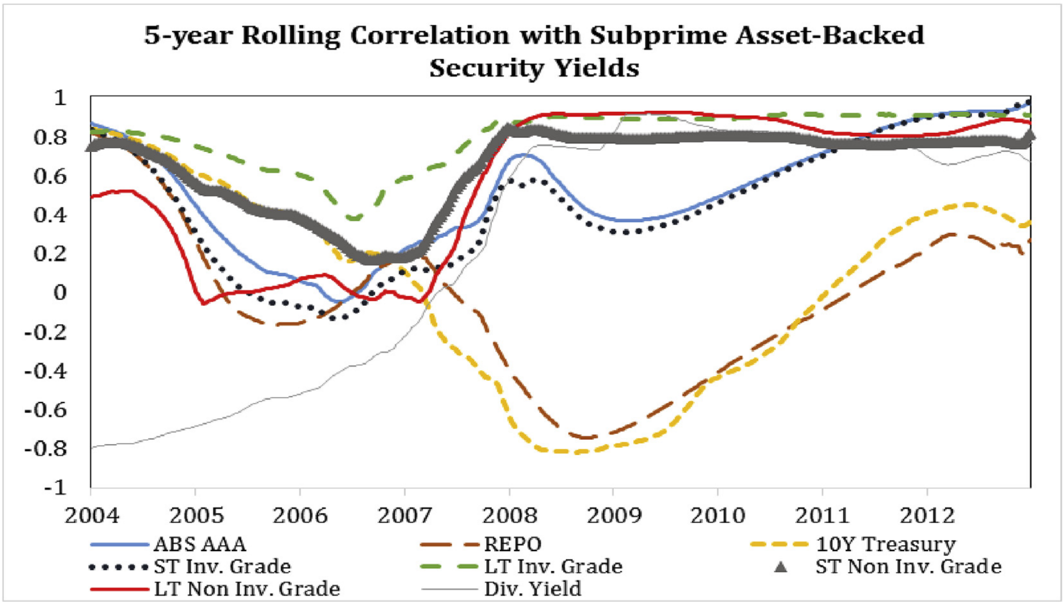

Fig. 1. Rolling window correlations with sub-prime ABS yields.

to quantify how a shock in the ABS 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 break 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 versus 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)[16]; 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 yields a set of eight different models.

The first column of Table 2 reports 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 in 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 favor comparisons, we compute 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 (ELL) of approximately -3031 is considerably lower than all other estimates in the table. Moreover, a MSVAR(q) model, in which U.S. financial markets simply shift between tranquil and crisis regimes does only slightly better, with an ELL of -2474. In Table 2, q is set to 1, as VAR frameworks with longer lags never provided a better in-sample fit.[17] On the one hand, this comparison confirms the result of previous MSVAR papers (see Guo et al., 2011; Guidolin and Pedio, 2017), which suggests that capturing regime shifts

[16] 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 $K_{1t} = 0$.
[17] Complete tabulated results are available from the authors upon request.

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. $\Pr(K_{It}=1)$ 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 $K_{2t}=0$.

	Exp. Log-Likelihood	Pr(K1t = 1)	Pr(K2t = 1)	Pr(K2t = 1)
Single-state VAR	−3031.0	____	____	____
MSVAR	−2473.8	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	−209.3	0.145	0.344	0.071
MI TVP VAR Homoskedastic	−198.2	0.146	____	____
MI TVP VAR Heteroskedastic	−245.2	0.000	0.358	0.058
CS-MI TVP VAR	−195.9	0.124	____	0.072

may be extremely important. On the other hand, it is natural to ask whether and by how much even more complex VAR models with TVP or mixture innovation may improve the fit.

The results in Table 2 suggest 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 −3031 to −657. In addition, moving from a MSVAR(1) to a SV TVP-VAR(1) improves the ELL from −2474 to −968. Interestingly, adding a SV component to a TVP VAR(1) reduces 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. 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 increases the ELL from −3031 (in the single-state VAR) to −198; a (purely) heteroskedastic model improves the ELL from −968 (in the SV TVP VAR model) to −245. Also in this case, the inclusion of heteroskedasticity fails to improve, per se, the fit of the model, as −245 falls below −198. Nonetheless the best fitting models are the complete, full mixture innovation model with ELL of −198 and especially the Cogley-Sargent modified mixture model with an ELL of −196. 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 $\mathbf{a}_t = \mathbf{a}_{t-1} + K_{2t}\zeta_t$) is modest at best. This represents a first, key result of our analysis: financial markets in the U.S. 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.

Finally, Table 2 reports implied break probabilities for the full and CG mixture TVP-VARs that are sensible. According to the full model, breaks in VAR, correlations, and variance coefficients have smoothed probabilities of 0.145, 0.344, and 0.071 (with the second estimate likely to be unreliable), respectively. This can be read as saying that episodes of crisis and contagion on average would last 7, 3, and 14 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 $\widehat{\Pr}\ (K_{1t}) = 0.164$ and $\widehat{\Pr}\ (K_{3t}) = 0.082$ and imply more plausible durations of 8 and 14 weeks.[18]

4.2. Patterns in parameter time variation

As previously discussed, according to a few mainstream definitions (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, what is sometimes defined as *shift-contagion* (Rigobon, 2016). 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, in this section, we take as given the posterior estimates obtained in Subsection 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.[19] However, given that it is impossible to report plots for 90 VAR coefficients, we report only the loadings of each yield on the low-grade ABS ones, which feature the instance of contagion that we want to study.

[18] Note that the plausibility of the break process cannot be simply assessed based on their frequency because even when relatively frequent, breaks follow a random walk process and when the variance of the shocks is modest, frequent breaks may be of modest size.

[19] 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).

4.2.1. Standard TVP VAR models

In Fig. 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 14th-86th percentile 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 2007–2009 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 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.[20] 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 vanished by the end of 2013.

Fig. 3 provides a similar comparison with reference to the posterior estimates of the own-shock volatility parameters for the three models in Fig. 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 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, Fig. 3 shows that all series (but the dividend and non-investment grade short-term yields) repeatedly spike up between 2007 and 2009.[21] 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 disentangle 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 Fig. 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 Fig. 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 GFC were purely induced by contagion and in no plausible way by simple increased interdependence.[22]

4.2.2. Mixture innovation TVP VAR models

Fig. 5 reports plots comparable to Fig. 2, but referring now to the two types of mixture innovation TVP-VAR models entertained in this paper (i.e., full model with no restrictions on $\mathbf{K}_t$ and the Cogley-Sargent variant with $K_{2t} = 0$). Even though Table 2 reveals some evidence in favor of a homoskedastic MI model over the full one, it remains pedagogically instructive to see how and where the CS restriction that fixes the correlations of the shocks to be constant affects the empirical results. As in Fig. 2, also in Fig. 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. Similar to Fig. 2, also in Fig. 5, we report 16th-84th percentile confidence bands, as is typical in the literature (see Primiceri, 2005) around the posteriors computed for the best fitting model (i.e., CS mixture model).

On the one hand, the results in Fig. 5 are less homogeneous and hence impressive when compared to the simpler, TVP-VAR cases. When the (wide) confidence intervals are taken into account, coefficients are subject to contagion exclusively during the financial crisis: coefficients turn significantly positive between August 2008 and June 2009. In particular, VAR coefficients of low-quality ABS, high-quality short-term corporate, and junk long-term corporate yield all increase significantly, which implies that shocks to low-quality ABS yields were transmitted with abnormal strength during the financial crisis (interestingly, starting between 1 and 5 months before Lehman Brothers's collapse in September 2008). The propagation of a downturn in the subprime ABS market to the investment grade market has been anecdotally discussed in much of the

[20] 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.

[21] 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 it is during 2008–2009 that their connection to other asset markets undergoes a deep change.

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

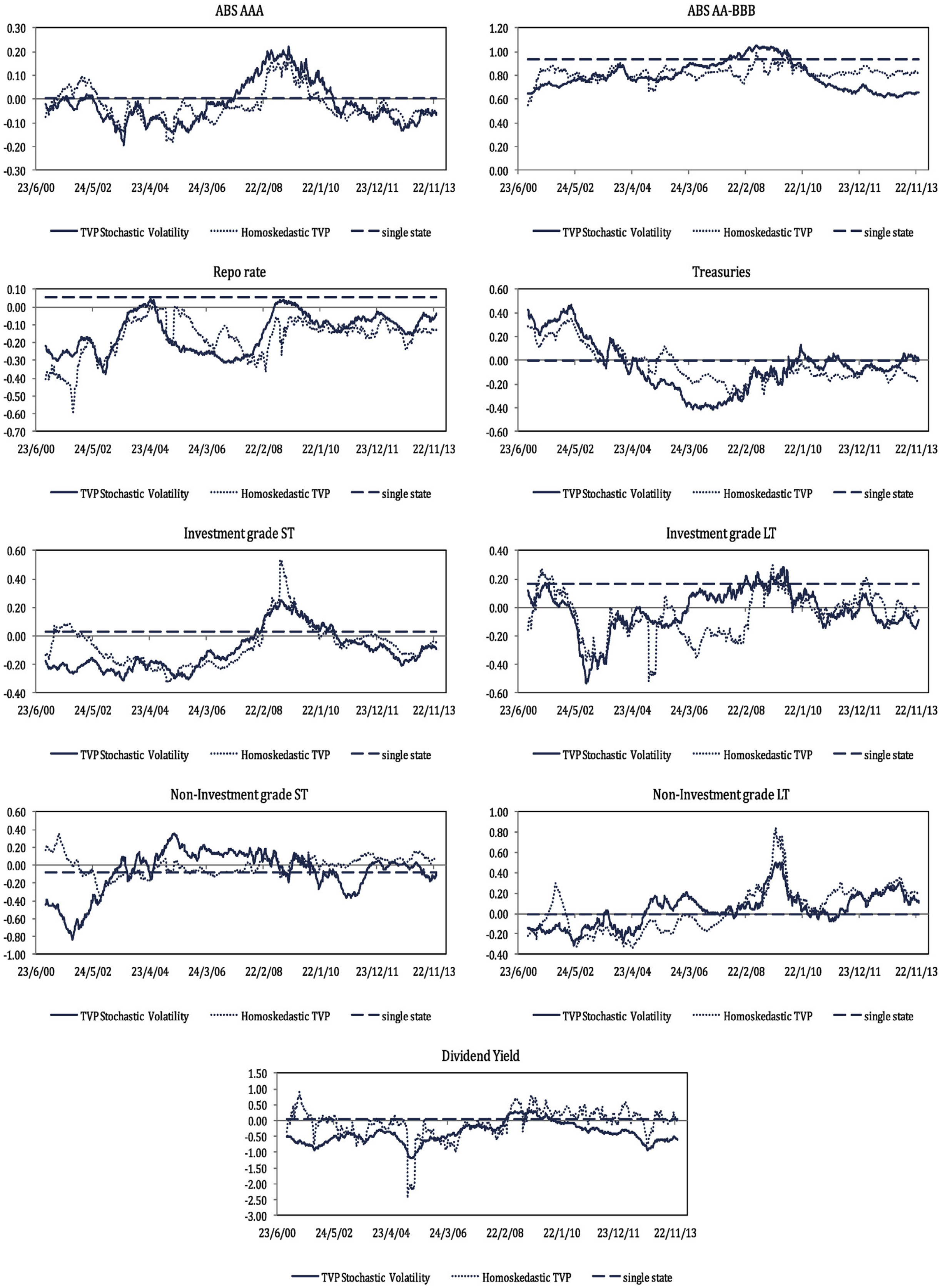

Fig. 2. Time variation of VAR coefficients in standard TVP VAR models.

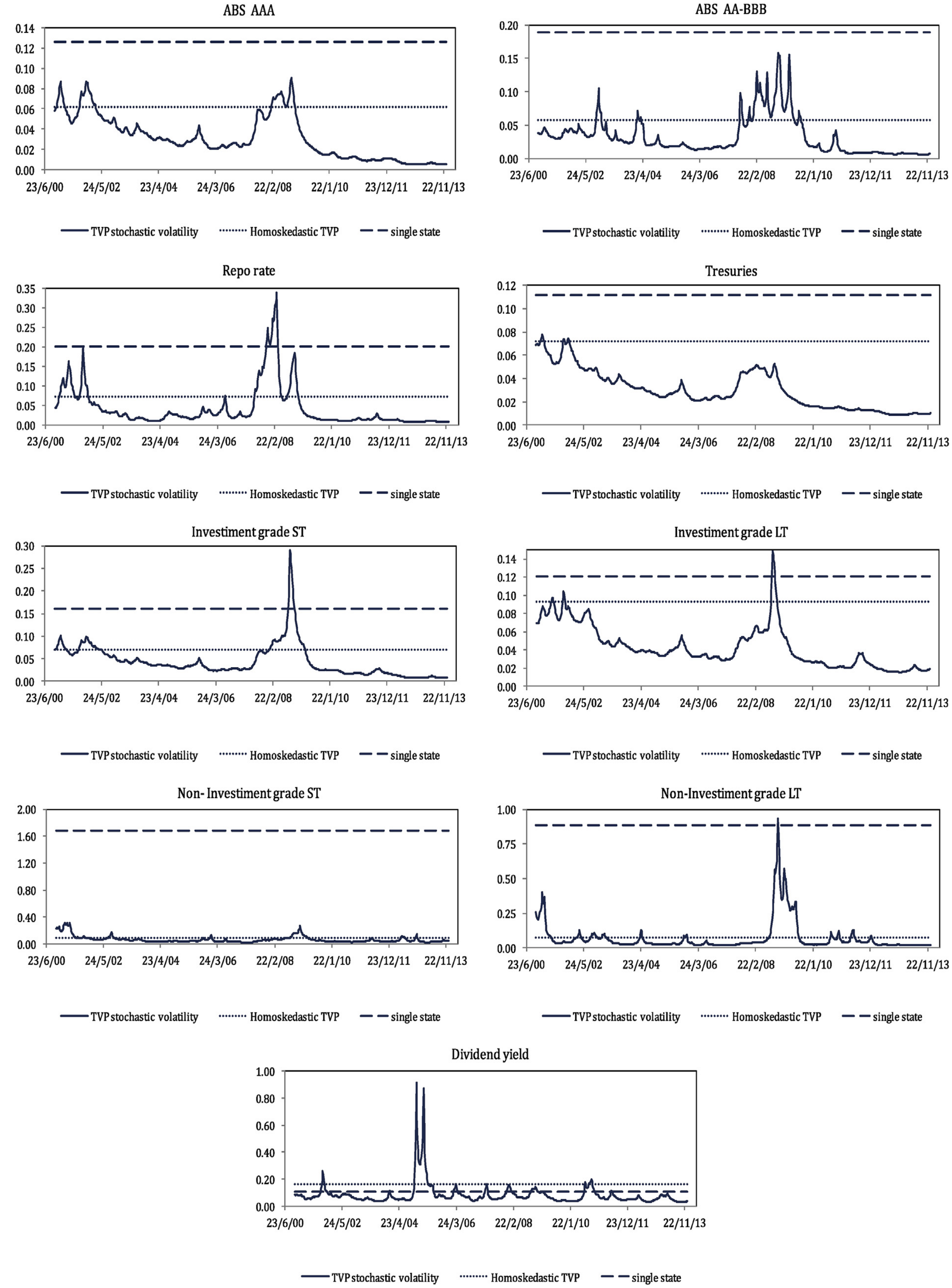

Fig. 3. Time variation of volatility coefficients in standard TVP VAR models.

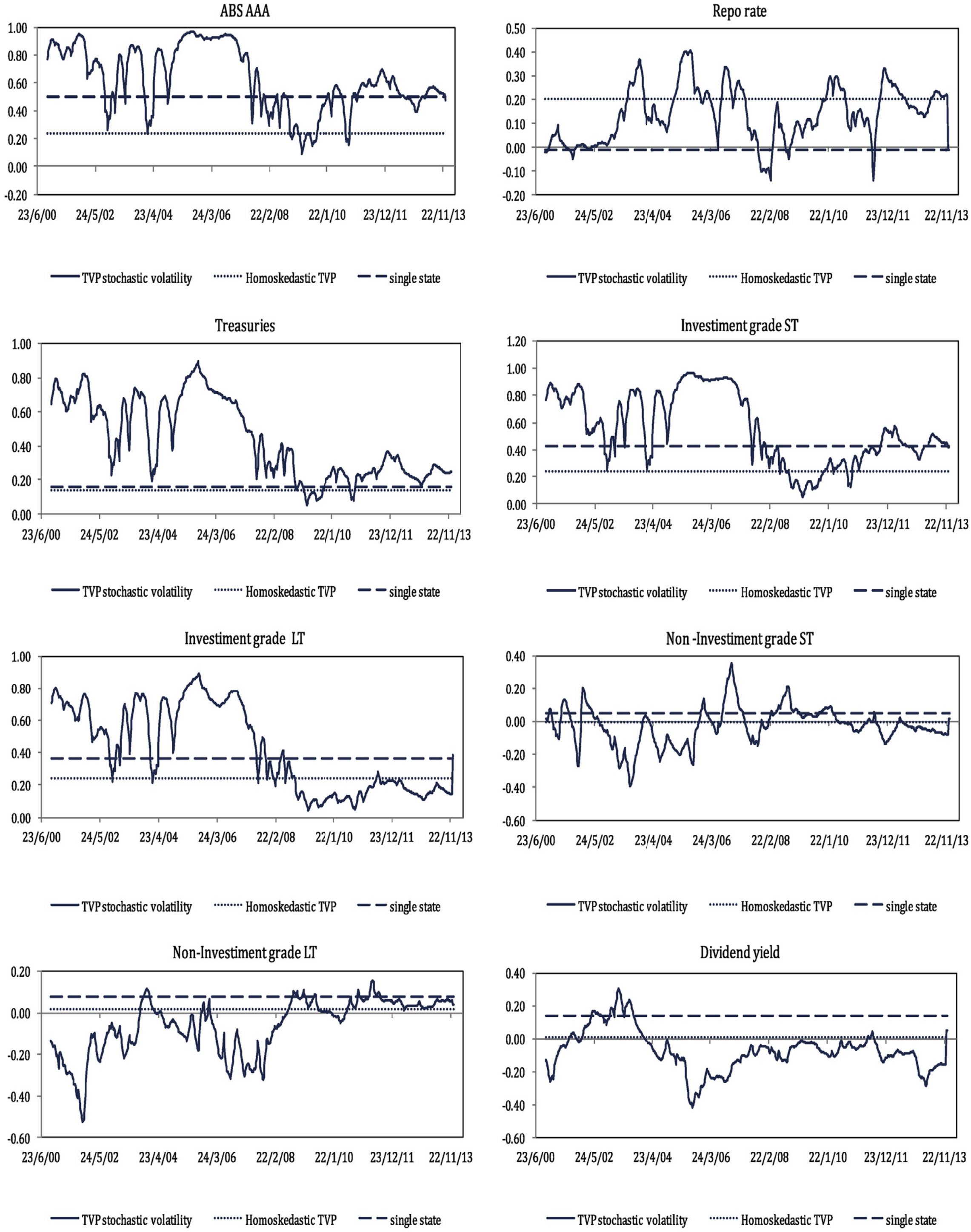

Fig. 4. Time variation of correlation coefficients in standard TVP VAR models.

literature on the GFC (e.g., Gorton, 2009). There is also a strong flight-to-quality effect, which turns out to be long-lasting in the case of Treasuries (2008–2012). Interestingly, it is not the case that the posterior means of the coefficients all move in homogeneous ways for all the risky assets: in the case of the long-term corporate yields of high-quality firms and of short-term junk yields, the effects are modest and—if any—they point towards yields being reduced by a shock to low-quality ABS. At least in the case of Aaa corporate bonds, this may reflect an additional flight-to-quality effect. Such a contemporaneous

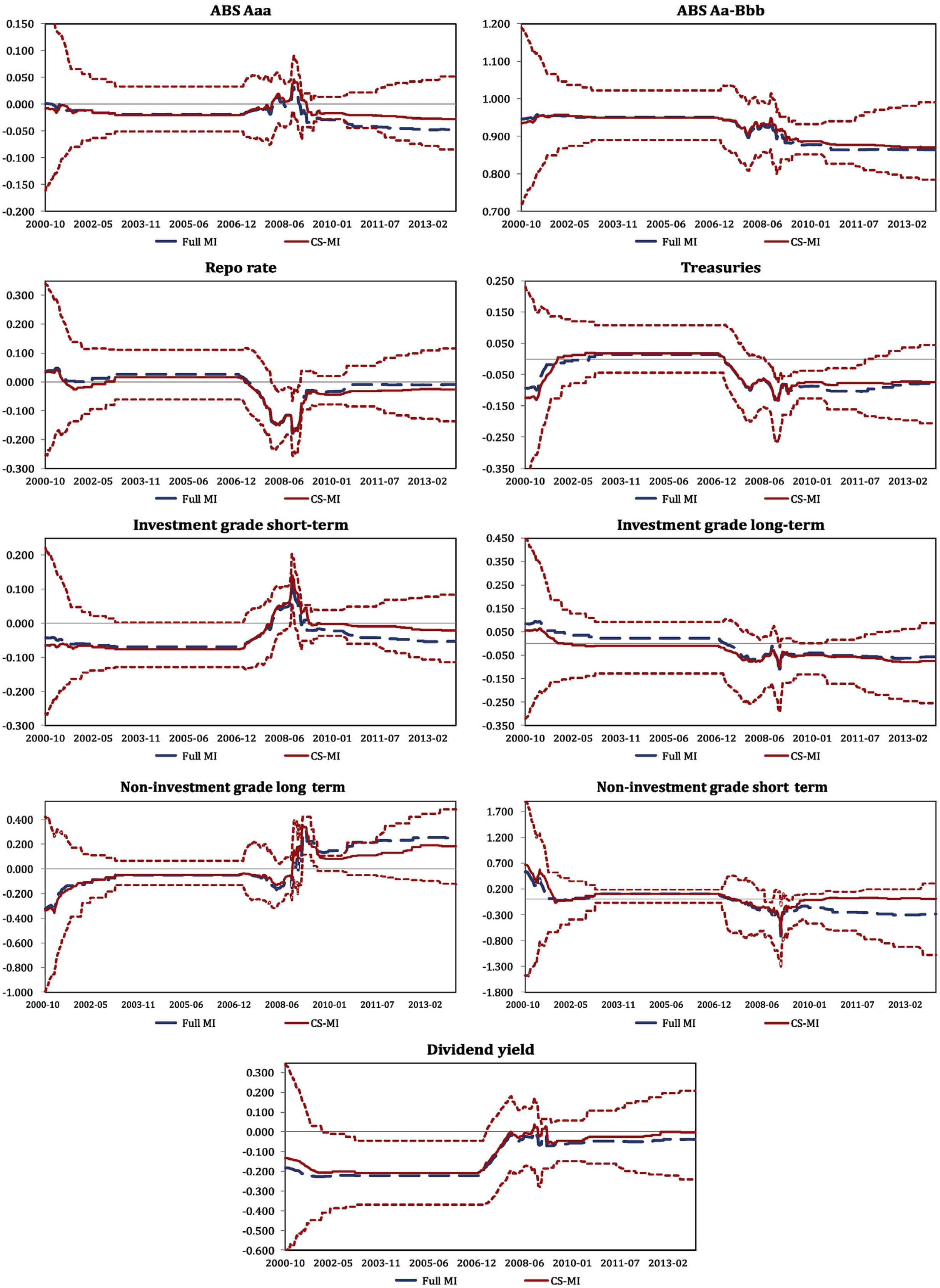

Fig. 5. Time variation in VAR coefficients in mixture innovation TVP VAR models.

contagion and flight-to-quality effects are exactly what we would expect in the light of the now ponderous literature on the crisis. The only mystery in Fig. 5 is posed by non-investment grade, short term corporate yields, which 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 subprime type ABS markets, investors react by moving their funds towards short-term junk bonds.[23]

It is of course important to also ask whether the dynamics in the loadings/VAR coefficients in Fig. 5 arise from sensible ex-post inferences on the location in time of breaks (as governed by K_{1t}) and their posterior, smoothed probabilities. Fig. 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, 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 0.27 of recording a conditional mean parameter break on every week; over the 2008–2009 sub-sample, such a probability is estimated at 0.18; over the remaining, non-crisis sample, the posterior frequency is instead slightly in excess of 0.01, as one would expect. This establishes that according to mixture TVP-VAR models, it is crisis samples that are dominated by breaks in coefficients.

In Fig. 7, we plot the posterior estimates of the time-varying volatility coefficient. The results are qualitatively similar to that in Fig. 3: according to all models that do not restrict K_{3t} to be zero (this is the case under the 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 (e.g., Yang and Zhou, 2016). Fig. 6, bottom panel, reflects (without recording the signs of the changes in volatility of 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, $\widehat{\Pr}_{t|T}(K_{3t}) = 0$.

Fig. 8 reports similar findings with reference to correlation coefficients between the shocks to each of the asset classes and low-grade ABS yields.[24] For most models and markets, such correlations appear to be either constant or declining—with Fig. 5 this delivers the picture of a U.S. financial markets in which dynamic linkages across markets are getting stronger and are subject to contagion effects, but in which unanticipated shocks become less interrelated. Even though there is little evidence of an increase in shock correlations corresponding to the U.S. financial crisis, this effect is weak and appears to strictly depend on the model investigated. In fact, Fig. 6 shows that the smoothed $\widehat{\Pr}_{t|T}(K_{2t})$ are generally constant, only with a minor uptick in 2008–2009. All in all, Figs. 5 and 8 show that during the financial crisis there was contagion of considerable strength from the low-quality ABS market to other risky segments of the U.S. 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 in the strength of the interdependence among markets (i.e., there is no evidence that correlated shocks hit all the markets simultaneously).

As previously discussed, in our definition of contagion two empirical features are critical: (i) the (posterior mean of) VAR coefficients describing the dynamic linkages among markets must shift over short intervals of time (heuristically, we may say, they are expected to jump) in directions that make markets more connected; (ii) such shifts/movements in coefficients must be accurately estimated. In Table 3, using the time series of the posterior VAR coefficient estimates of the loadings connecting yields to past lower-quality ABS yields, we present a few statistics pertaining to these two aspects. Of course, treating the posterior means of coefficients as if they were data poses some issues, but we propose the statistics in Table 3 just to gain additional insight and without formal inferential objectives. In columns (1)–(4) of the table, we report the posterior means of the coefficients loading on each of the eight series under analysis on lower-rating ABS yields in the pre- (Oct. 2000–June 2007) versus the post-crisis (July 2007–June 2009) periods and a few related statistics. With one exception already discussed above (non investment grade, short-term corporate yields), all other assets record massive shifts in the posterior means of the VAR coefficients in the expected directions: the mean loadings of risky yields all increase by percentages ranging between 86% (equities) and 2387% (investment grade corporate, long-term); the mean loadings of repo contracts and Treasuries decline by 1202% and more than 111,000% (such a high value occurs as Treasury rates were unaffected by low-grade ABS yields in the pre-crisis period). These figures leave little doubt on the entity of contagion between late 2007 and 2009. In columns (5)–(7) of Table 3, we report the strongest reactions of the posterior means of VAR loading coefficients and its speed, in terms of number of weeks between the peak (for risky assets) and the trough (for Treasury and repo rates) and the pre-crisis average. All such differences between peaks/troughs and the pre-crisis average are large and precisely estimated. In terms

[23] In contrast with the traditional literature where contagion is defined as an increase in cross-markets linkages, Corsetti et al. (2011) show that, depending on the nature of the crisis, the related parameter values could either decrease or increase during periods of market turbulence. In line with their findings, our cross-asset relationships do not always strengthen in distressed periods. However, we remain surprised by our empirical findings.

[24] Correlations are (modestly) time-varying in the CS TVP VAR framework because, even though $K_{2t} = 0$, K_{3t} remains time-varying and (for given covariance) the variation of the standard deviations is sufficient to induce time variation in correlations.

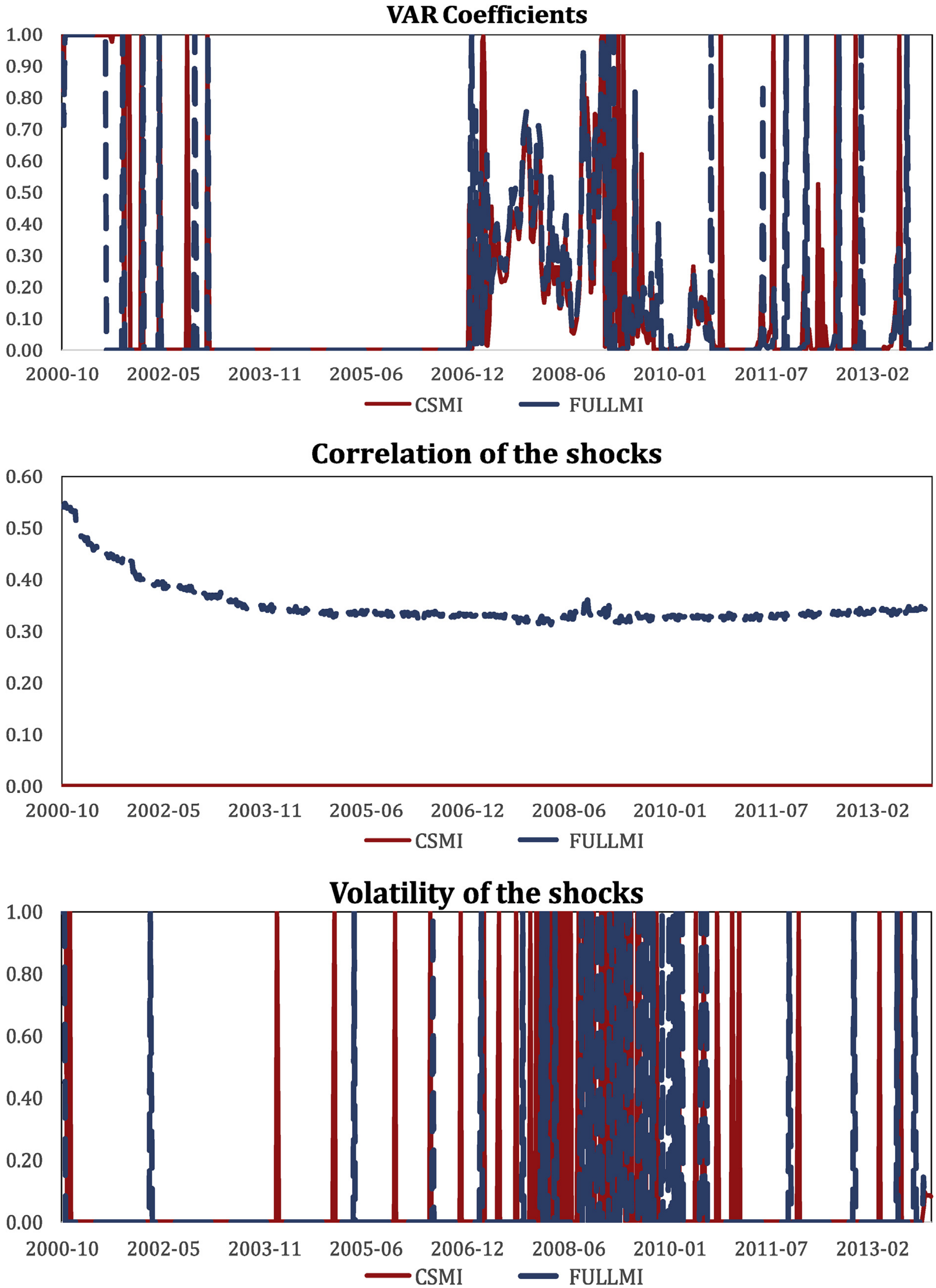

Fig. 6. Posterior estimates of the probabilities of breaks from mixture innovation TVP VAR models.

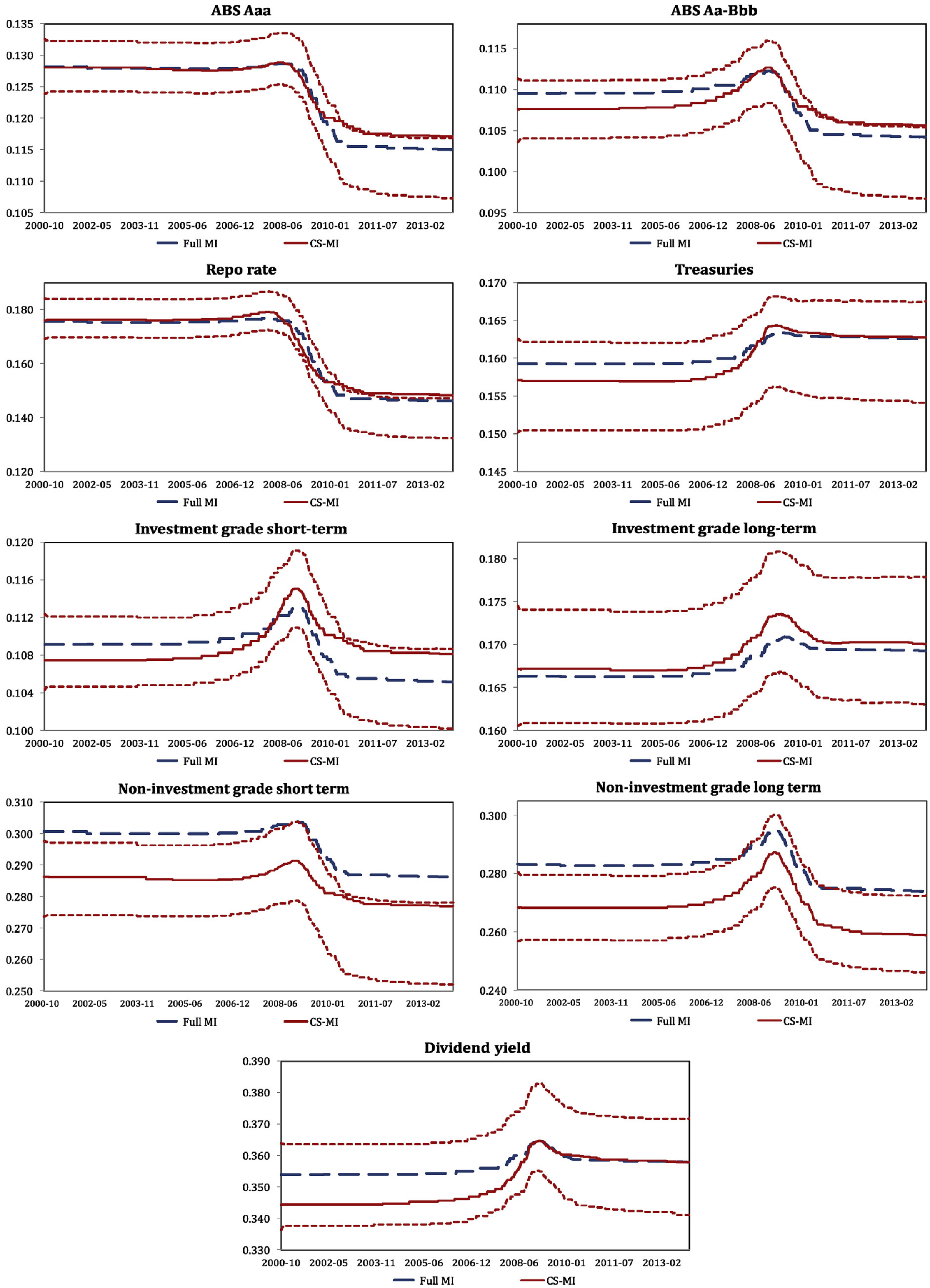

Fig. 7. Variation of volatility coefficients in mixture innovation TVP VAR models.

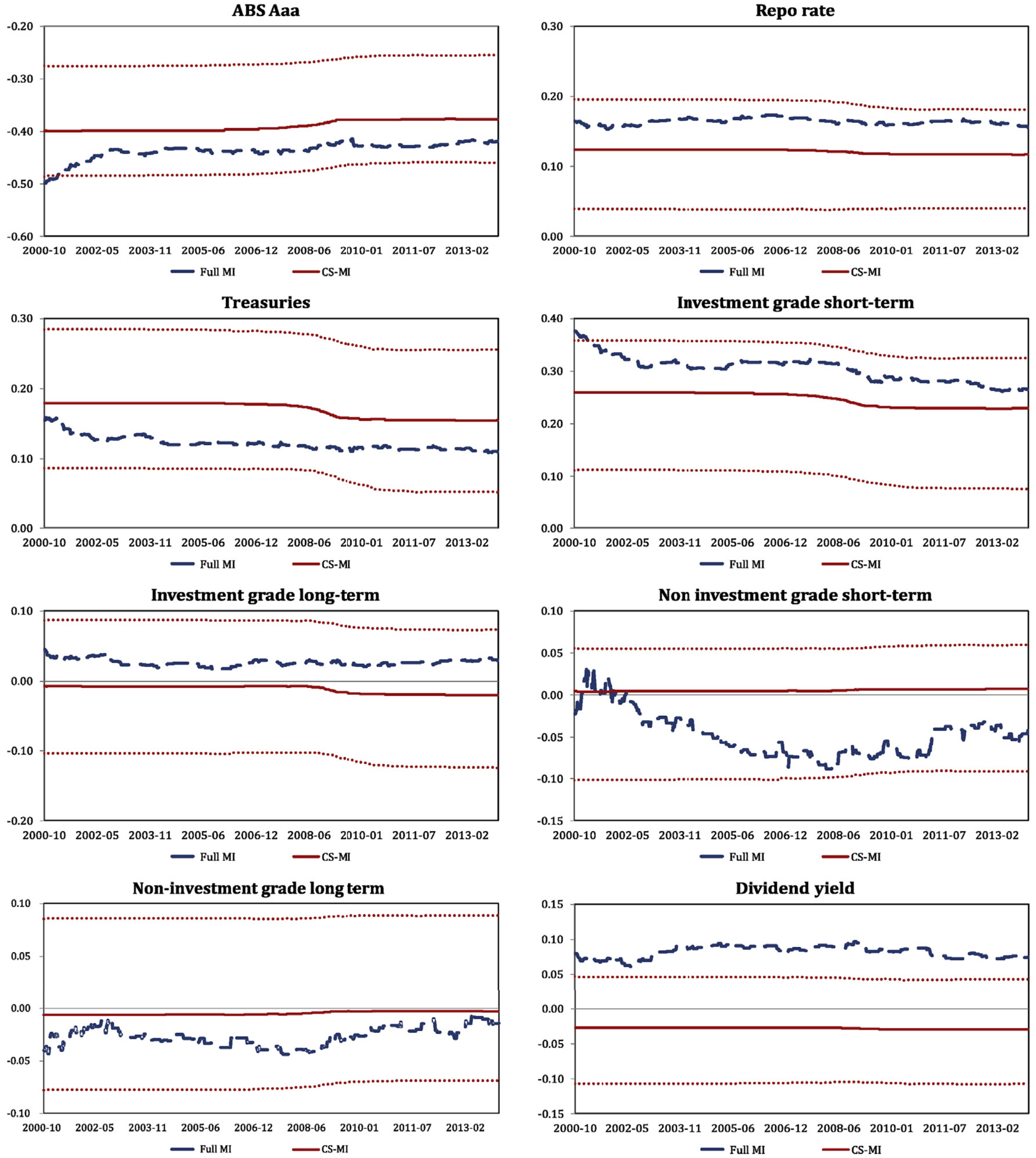

Fig. 8. Time variation of correlation coefficients with ABS Aa-Bbb yields in mixture innovation TVP VAR models.

of speed, such peaks and troughs are all reached in a number of weeks that ranges between 20 (equities) and 84 (non investment grade, long-term corporate rates); for most series, the number of weeks required for the strongest effects to appear is around 70 weeks and peaks/troughs are all attained between September 2008 and February 2009 (the only exception is the dividend yield, characterized by a peak of sensititivity to low-grade ABS in 2007). While 70 weeks are just around 10% of the overall sample size and well agree with the visual indications provided by Fig. 5, note that counting from the first week of July 2007 remains rather arbitrary. If we were to date the beginning of the U.S. financial crisis in correspondence to the first week of August 2008 (see Guidolin and Tam, 2013, for related evidence and a review of similar dating exercises), then such a speed statistic would decline to about 15 weeks, which seems to indicate a fast reaction.

Table 3
Average Posterior Mean Coefficient Changes from Cogley-Sargent's Mixture TVP VAR Model and Speed of Reaction. The table compares the posterior means of the coefficients loading each series on ABS Aa-Bbb yields in the pre- (Oct. 2000–June 2007) and post-crisis (July 2007–June 2009) periods and computes a few related statistics. Moreover, in the table we report the largest absolute value of the difference between the peak/trough in the crisis period and the pre-crisis average and test for its statistical significance using the sample standard deviation of the differences between the crisis and pre-crisis posterior means of the coefficients. The asterisks attached to the statistics mean *** significant at 1%; ** significant at 5%; * significant at 10%.

	Average 2000–2006	Average 2007–2009	D due to crisis	Relative % Change	Peak to average	Date of Peak/Trough	Number of Weeks since July 2007	Notes
ABS Aaa	−0.017	0.004	0.022	126.1	0.063***	Oct. 2008, third week	68	
GC Treasury Repos	0.009	−0.097	−0.106	1201.5	−0.186***	Oct. 2008, third week	68	
10Y Treasuries	−0.000	−0.083	−0.082	111,603.8	−0.133***	Dec. 2008, third week	76	
Investment grade corporate, short-term	−0.073	0.027	0.100	136.9	0.213***	Sept. 2008, third week	64	
Investment grade corporate, long-term	0.002	0.060	0.057	2386.6	0.101**	Dec. 2008, fourth week	77	
Non investment grade corporate, short-term	0.125	−0.153	−0.278	−222.8	−0.819***	Jan. 2009, second week	79	Unexpected sign
Non investment grade corporate, long-term	−0.085	0.026	0.111	130.8	0.427***	Feb. 2009, second week	84	
Dividend yield	−0.197	−0.027	0.170	86.3	0.234***	Nov. 2007, third week	20	

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 we investigate. 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.[25]

4.3.1. *Standard TVP VAR models*

The time-varying nature of the parameters of all of these models implies that 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 Fig. 9, we plot the IRFs for a homoskedastic TVP-VAR model in correspondence of four selected dates: mid-October 2000 and mid-September 2008 (to capture the effects of the dot-com crisis and of the sub-prime crisis, respectively), early January 2005 and the last week of December 2013 (to describe how the transmission mechanisms settled down some years after the end of each of the two crises). Similarly, in Fig. 10, we plot the IRFs for a TVP-VAR model with stochastic volatility at the same dates.[26] 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, the IRFs corresponding to mid-September show a different behavior: especially for the riskier assets, they are much larger and, most importantly, they systematically increase over a 21-week horizon (but they stabilize and eventually decline, as they should given that stationarity of the models has been enforced during the estimation process).[27] Crucially, all the IRFs jump up and imply much stronger effects (between 4 and 20 times the impact recorded outside crisis periods), which signal the existence of contagion effects. A comparison of Figs. 9 and 10 reveals that including stochastic volatility improves the ability of TVP-VAR models to detect episodes of contagion. More specifically, the breaks that alter the position and shape of the IRFs are larger in Fig. 10 than in Fig. 9, 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.

[25] We omit confidence intervals from 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.

[26] We decided to report the IRFs at four specified dates to enhance readability. However, the full, three-dimensional plots of IRFs for the homoskedastic TVP-VAR model and the TVP-VAR model with stochastic volatility are available in an Online Appendix (Figs. A3 and A4, respectively).

[27] Also in Figs. 9 and 10, the effect of a low-grade ABS shock is negative on repo and Treasury yields. Interestingly, the dot-com crisis caused very different reactions when compared to the recent financial crisis. In particular, during the dot-com collapse, 10-year Treasury rates display a non-monotonic, positive response, the opposite of the negative, immediate reaction that characterizes the subprime crisis.

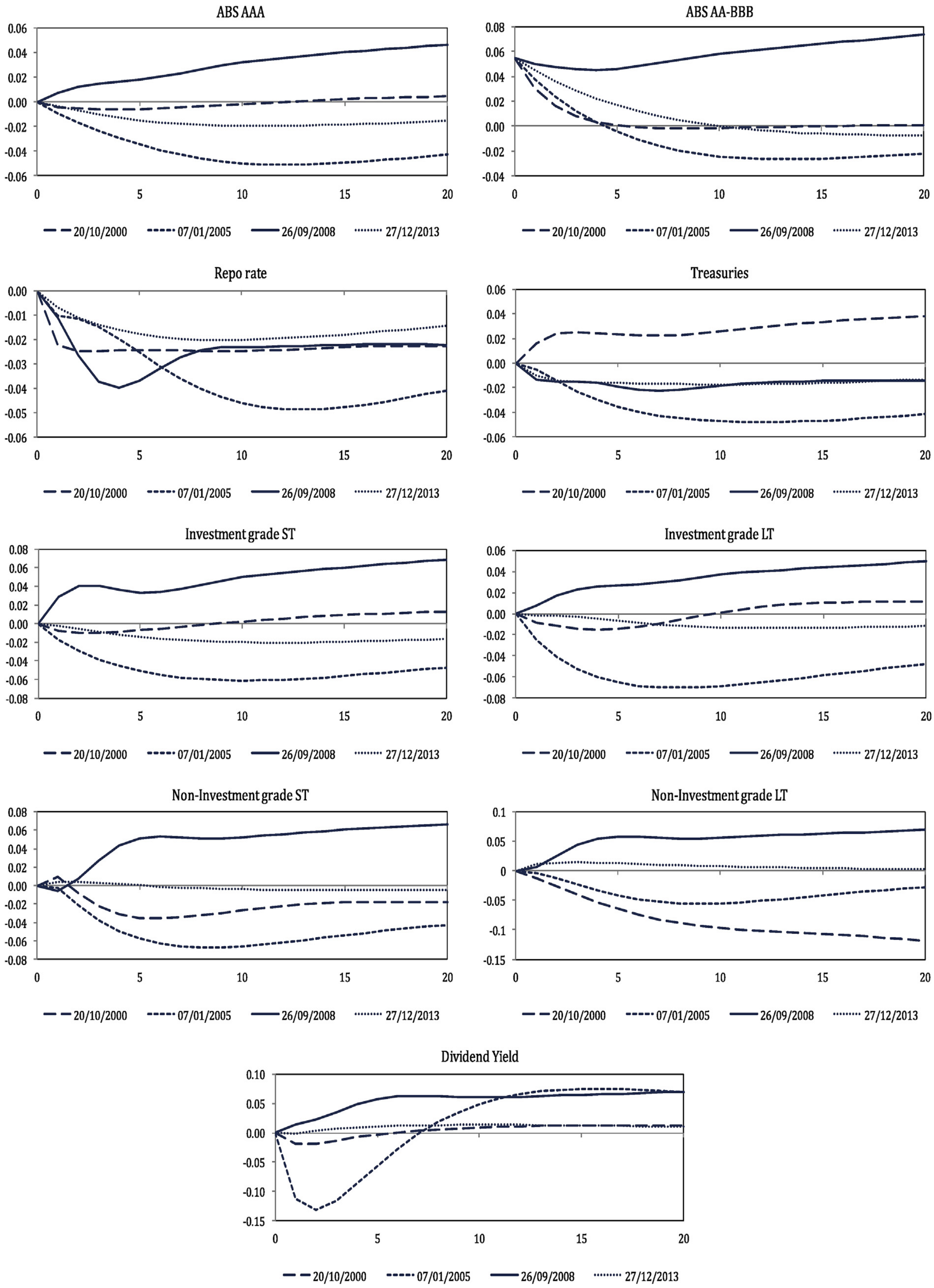

Fig. 9. Selected IRFs implied by a homoskedastic TVP VAR model.

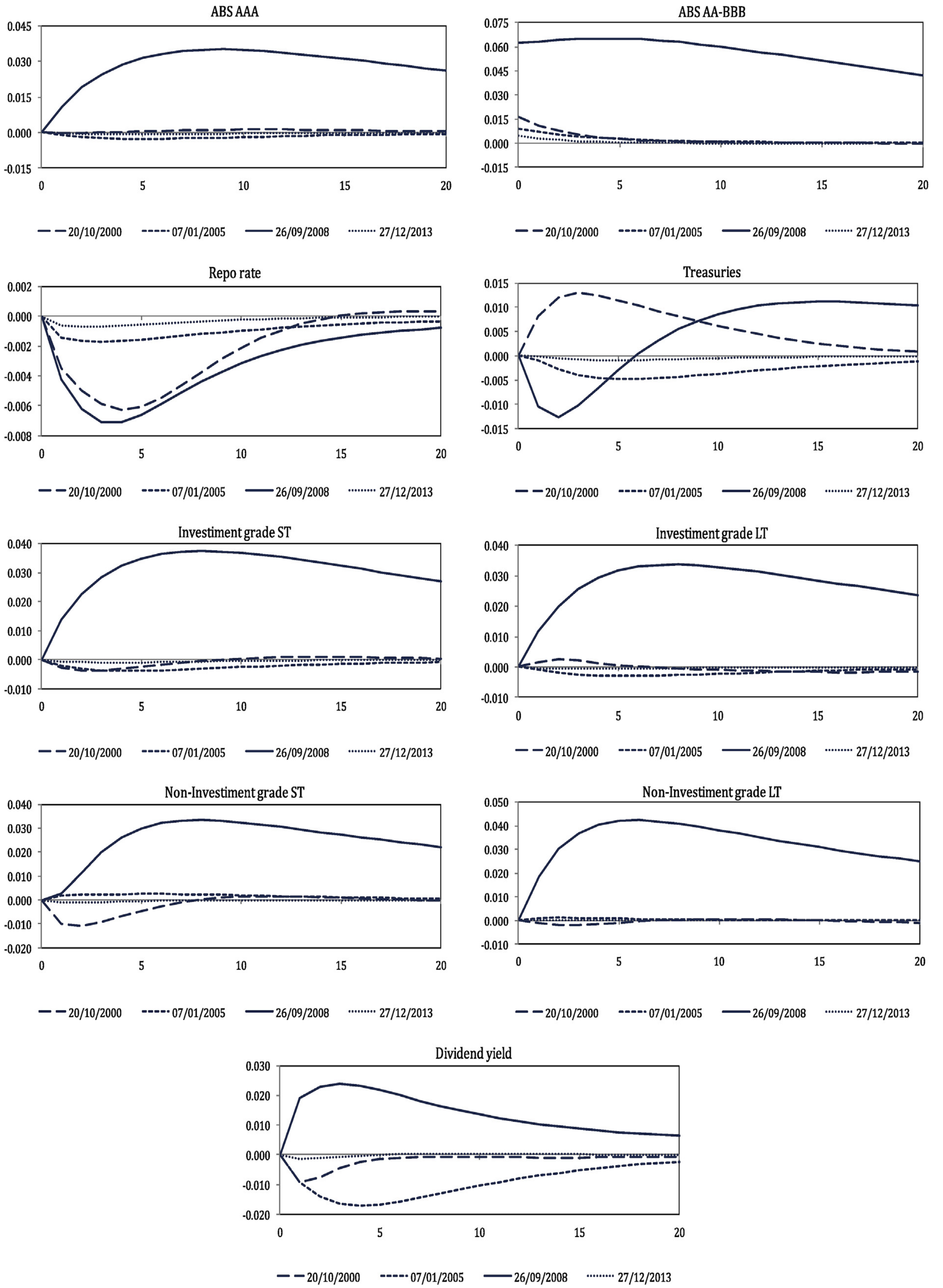

Fig. 10. Selected IRFs implied by a TVP VAR model with stochastic volatility.

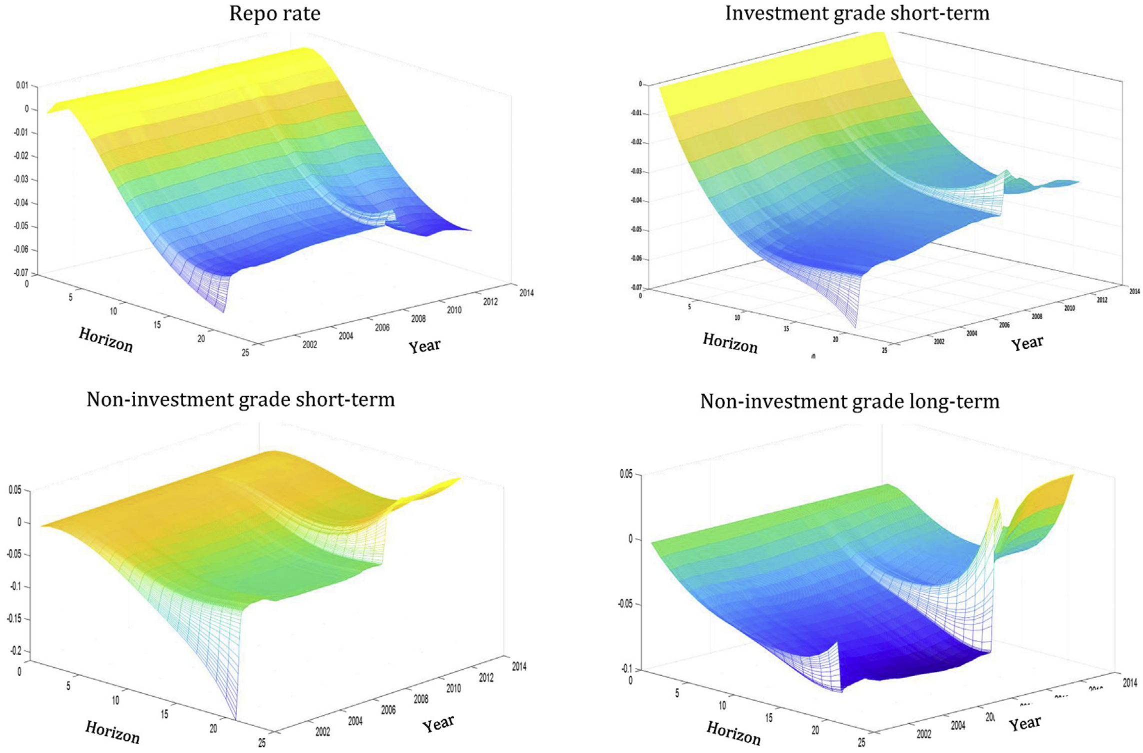

Fig. 11. IRF by the CS mixture innovation TVP VAR −shock to a low-grade ABS yield.

4.3.2. Mixture innovation TVP VAR models

Due to their similarities, we present the results from the mixture innovation TVP-VAR models only with reference to the CG model. The results for the full mixture framework were visibly indistinguishable for our purposes. Figs. 11 and 12 show the three-dimensional and the bidimensional plots (for the same selected dates discussed in the previous Subsection), respectively. In general, despite the introduction of the mixture innovation variables, the overall pattern of the IRFs is similar to those reported for the simpler TVP-VAR models in Figs. 9 and 10. However, there are few exceptions. In particular, comparing Fig. 11 and A3-A4 (in Appendix), 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 identify, a shock to lower-grade ABS yields generates weaker and smoother reactions in other series, especially in corporate rates, when compared to the case in which VAR parameters are forced to undergo continuous changes in Fig. 10. These results are similar those discussed in Subsection 4.1, where contagion is starker when time-variation in parameters is 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 our sample period. 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 drive the effects observed during the financial crisis. The analysis is carried out for the CG version of the mixture innovation TVP-VAR, but the results are not sensitive to adopting other frameworks 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 investors' 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 examine.[28] Moreover, the demand for Treasuries is also reflected by the

[28] 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.

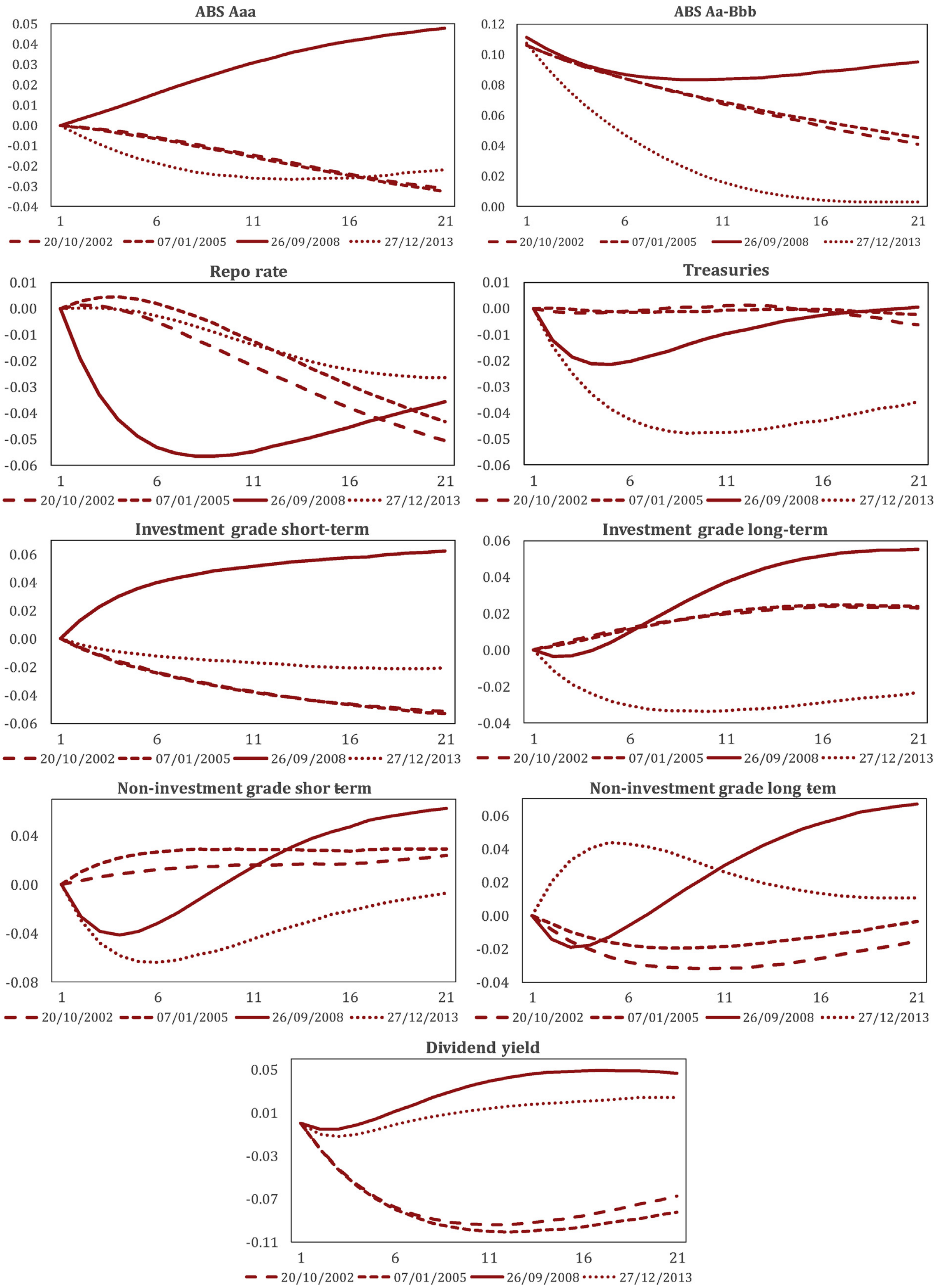

Fig. 12. Selected IRFs implied by the CS mixture innovation TVP VAR model.

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.

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 period in which these effects turn weaker (but still persist, see Fig. 11) is 2003–2006, when the corresponding IRFs are flat. 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 throughout the entire sample. Interestingly, this effect strongly intensified from the outbreak of the subprime crisis until 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 (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 2007–2009 GFC in the U.S. This increase in risk premia is explained by what he refers to as 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 dynamic 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 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 report a simultaneous and precisely estimated depreciation of only some of these risky yield series and this does not lend strong support to the uniform presence of a risk premium channel, while the combination of positive or modest 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.

Both short- and long-term non-investment grade corporate bonds display a positive cumulative response to shocks to lower quality ABS yields during the subprime crisis. However, in Section 4 a difference has 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, during the recent financial crisis, the risk premium channel is also at work. These findings could be explained by the extremely widespread uncertainty that characterize the U.S. 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 investment grade corporate bonds reveal instead heterogeneous reactions depending on whether we consider short- or long-term paper: yields on long-term corporates remarkably decrease while the opposite occurs to short-term ones. These results suggest that risk premium channels affected only short-term, riskless corporate paper. An explanation of this finding is that, after the negative shock to the ABS Aa-Bbb, the worst (marginal) companies included in the investment-grade corporate bond portfolio were quickly downgraded and the credit level of the overall class increased, thus becoming a safer asset class, which is however fully consistent only with a flight-to-quality channel (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 into the prices of other assets, through a learning mechanism (e.g., Kodres and Pritsker, 2002). 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 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 high yield increase characterizing the former (15 bps) can be almost totally explained by a correlated information channel.

6. Robustness checks

In this section, we describe a few additional experiments that we have conducted using our econometric framework. In qualitative terms, the empirical findings reported above tend to be fairly robust, even though the exact quantitative estimates and impact assessments depend on the details of the experimental design adopted.

Up to this point, in the case of the best fitting mixture innovation TVP VAR models, we have worked under fairly skeptical priors concerning the *per-period* probability of a break, which was set to be approximately 1.06% per week. This seems sensible because in a formal sense, our statistical null hypothesis consists of the absence of contagion and therefore of breaks/regimes in the VAR coefficients of the model. Equivalently, as discussed in Section 3, we did aim at persuading the data to reveal—if any—the existence of contagion. Even though such a small per-period probability implies that the cumulative probability of at least one break over our 14-year sample is in excess of 99%, this may seem restrictive and poses some risks. For instance, if such a prior were to allow a few (in the limit, one) structural shift episode to emerge from the data, it would provide unwarranted special emphasis and prominence to the 2007–2009 subprime/financial crisis episode and confound our chances to actually reject the null of the absence of contagion.

Therefore, as a first step, we double the prior implied probability (formally, we repet our entire analysis assuming $\underline{\beta}_{1j} = 0.1$ and $\underline{\beta}_{2j} = 50$ in the beta prior for the vector of breaks $\mathbf{K}_t$) of a break in each single period to 0.2%. This implies that over the entire sample of 656 weeks, one break now occurs almost surely and that several breaks are possible (e.g., a cumulative total of 6 breaks occurs now with a probability of 0.5). Of course, these are just prior probabilities: through the likelihood, especially with 656 observations, there are plenty of chances for the posteriors probabilities of a per-period break (see Table 2) to turn out to largely exceed such values of 1.1% or 0.2%. Complete tables and figures concerning this alternative parameterization of the priors are reported in an Online Appendix. Just to provide a feeling for what the results would look like, Fig. 13 parallels Fig. 5 and shows the dynamics of the posterior means for the VAR coefficients over our sample.[29] Qualitatively, there is no change in the results, although three comments are in order. First, as one should expect, the breaks are now more pronounced and frequent, even though the spikes and troughs revealed by the subprime crisis do remain unchallenged and their nature and entity hardly changes our comments above.[30] Second, a trade-off appears: while the step-like dynamics in the posterior of the coefficients typical of mixture model becomes more visible, the left-most scale of most plots shrinks, which indicates that compared to Fig. 5, all coefficients are pulled towards zero—while breaks are visible and relatively frequent, their simultaneous patterns tend to surrogate the otherwise estimable dynamic linkages among the series. Third, priors that are more liberal towards breaks would induce to a discussion of the existence of a contagion episode in 2000–2001, at least in terms of posterior means of the VAR coefficients, in the sense that for a few of the series (in Fig. 13 this is visible in the case of ABS Aaa and repo rates, and the dividend yield), these appear to be altered, different from the otherwise prevailing values, and in general to fall half-the way between the 2002–2006 estimates and those reached in 2007–2009. However, simple inspection of the corresponding confidence region would make such an analysis difficult, because with so few observations, the estimated posteriors have rather fat tails and systematically contain a value of zero for the coefficients. Interestingly, under this alternative set of priors, an analog of Table 2 shows that the full MI TVP VAR and its Cogley-Sargent's version deliver the same in-sample fit, but the former still has a hard time identifying the occurrence of breaks in the correlations of the residuals.

Alternatively, we have also entertained the case in which in our beta priors concerning the variables in $\mathbf{K}_t$, we set $\underline{\beta}_{1j} = 0.1$ and $\underline{\beta}_{2j} = 0.2$, to yield a very break detecting-oriented prior of 50% per week. Interestingly, such a prior implies that the probability of finding a break in exactly half of the weeks in our sample is again 50%. However, at least in qualitative terms, the difference between such a prior and the limit case of a fully heteroskedastic, TVP VAR model with SV as in (12)-(14) in which breaks are expected to occur on every single period, is modest at best. In fact, when we have performed full-information estimation, the results are almost identical to those shown in Figs. 2–4 and 10, for the case of stochastic volatility: breaks are of small size but they occur roughly in 70% of the sample in the case of VAR coefficients and in an even higher proportion

[29] In this case, we have omitted confidence regions around the posterior mean estimates, but they are available upon request. In general, their size increases moderately, which means that imposing a prior of more breaks extracts less precision from our data, as one should expect.

[30] This is confirmed by the analog of Fig. 6 that shows a flurry of weak (with posterior probabilities always below 0.3) break signals for the VAR coefficients in 2000–2001 and in 2006.

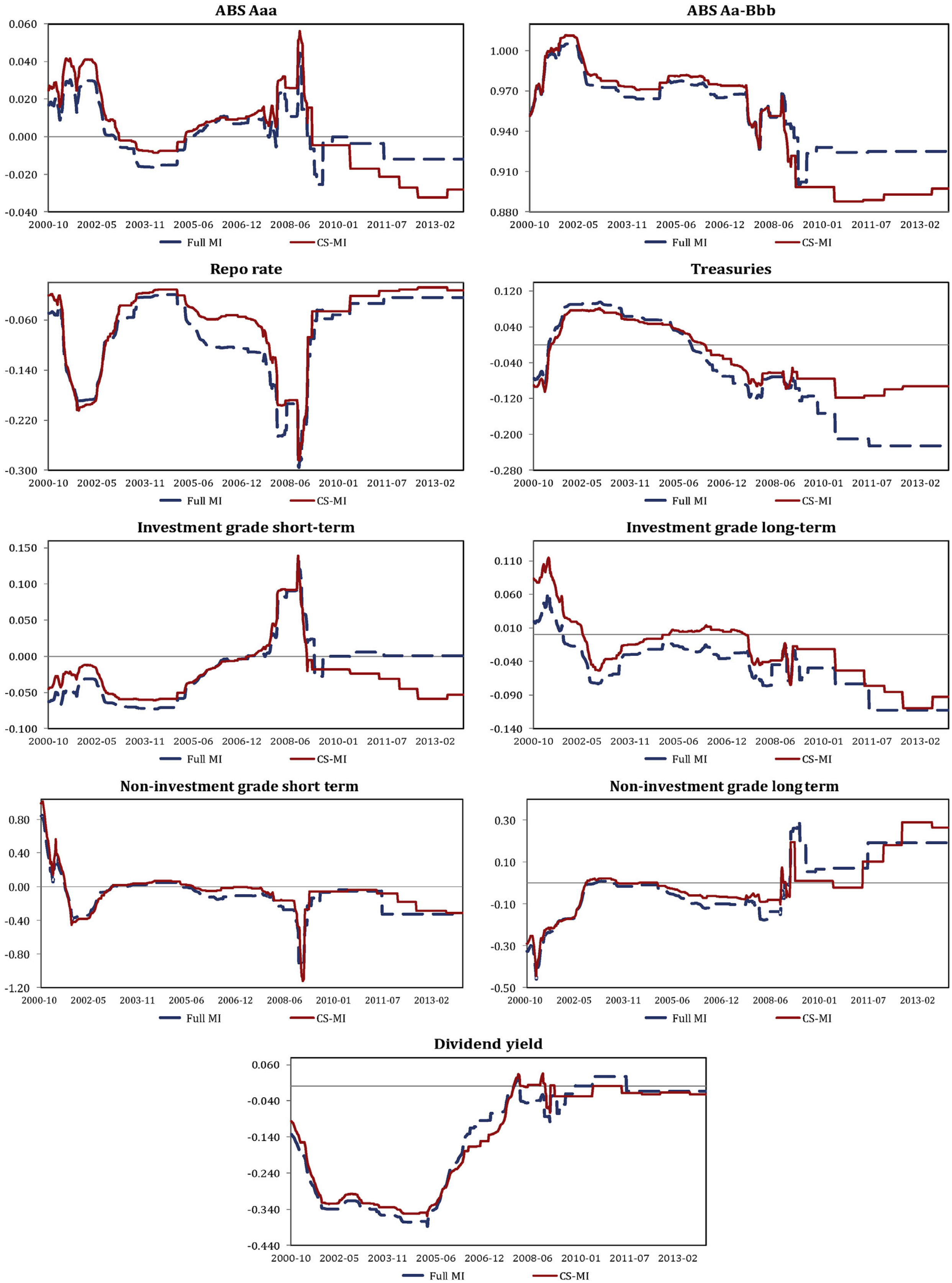

Fig. 13. VAR coefficients in mixture innovation TVP VAR models under alternative priors on the per-period probability of breaks.

of the weeks in the case of the volatilities and correlations of the shocks.[31] Finally, it should be clear that further experiments based on beta priors with $\underline{\beta}_{1j} = 0.1$ and $\underline{\beta}_{2j} < 0.2$ imply posterior distributions for the coefficients that are virtually identical to those in Figs. 2–4 and 11 and therefore are omitted to save space.

7. Conclusion

We investigate the occurrence of cross-asset contagion by modelling time-varying conditional mean parameters and heteroskedasticity in U.S. ex-ante yields. We study whether and when the transmission mechanisms of shocks across different markets have changed over 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 literature (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 obtain under the mixture innovation TVP-VAR model suggest that shifts in conditional mean coefficients intensely occurred at the outset of the subprime crisis and to some extent also at the end of the dot-com bubble. In contrast, 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 recent financial crisis: our GIRF analysis suggests that the transmission mechanism of shocks has changed over our sample period, revealing sustained interconnections pointing to contagion effects since late 2007, when the U.S. 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 2007–2009 turmoil. A detailed analysis of estimated IRFs reveals that the flight-to-liquidity as well as the flight-to-quality (and to some extent, the correlated information) channels were at work when the subprime crisis spread over the U.S. financial system.

There are a number of directions in which our work could be extended. Even though our goal was to show that thanks to MCMC Bayesian methods it is possible to estimate complex but flexible models that may allow us to disentangle contagion from spillover effects in volatility and correlations, we must acknowledge the existence of a rising literature that uses graph theory and statistical methods to estimate networks to perform inference on the structure of the interrelationships that represent the background of contagion dynamics (see e.g., Jackson, 2014). For instance, Cohen and Frazzini (2008) and Menzly and Ozbas (2010) find evidence of return predictability (necessary to any anticipated contagion) across economically linked firms and motivate it on the basis of the presence of investors subject to attention constraints by which asset prices would not promptly incorporate news about related firms. Sharifkhani and Simutin (2016) examine the same issues adopting explicit techniques from graph theory and show that trade network feedbacks exist, i.e., sectorial shocks propagating along a network can echo back to the originating industry via its linkages with trade partners, thus inducing intermediate-term autocorrelations in quantities of interest. Kelly, Lustig, and Van Nieuwerburgh (2013) provide a network-based explanation of stock volatility. In fact, recent work has been performed to show that the technology of Bayesian VAR models with regimes can be extended to estimate the structure of networks and this may provide and important point of contact between the two literature (see, e.g., Carvalho and West, 2007; Chong and Kluppelberg, 2018), with some early examples of applications to contagion becoming available, as in Yang and Zhou (2013).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.finmar.2019.04.001.

References

Abbate, A., Eickmeier, S., Lemke, W., Marcellino, M., 2016. The changing international transmission of financial shocks: evidence from a classical time-varying FAVAR. J. Money Credit Bank. 48 (4), 573–601.
Ang, A., Bekaert, G., 2007. Stock return predictability: is it there? Rev. Financ. Stud. 20 (3), 651–707.
Arakelian, V., Dellaportas, P., 2012. Contagion determination via copula and volatility threshold models. Quant. Finance 12 (2), 295–310.
Bacchiocchi, E., 2017. On the identification of interdependence and contagion of financial crises. Oxf. Bull. Econ. Stat. 79 (6), 1148–1175.
Bai, J., Julliard, C., Yuan, K., 2012. Eurozone Sovereign Bond Crisis: Liquidity or Fundamental Contagion. Working paper. Federal Reserve Bank of, New York.
Banerjee, S., Graveline, J.J., 2013. The cost of short-selling liquid securities. J. Financ. 68 (2), 637–664.
Bianchi, D., Guidolin, M., Ravazzolo, F., 2017. Macroeconomic factors strike back: a Bayesian change-point model of time-varying risk exposures and premia in the US cross-section. J. Bus. Econ. Stat. 35 (1), 110–129.
Billio, M., Caporin, M., 2010. Market linkages, variance spillovers, and correlation stability: empirical evidence of financial contagion. Comput. Stat. Data Anal. 54 (11), 2443–2458.

[31] The results are available upon request but when compared to the TVP VAR(1)-SV in Figs. 2–4, the differences were often difficult to detect.

Brenner, M., Pasquariello, P., Subrahmanyam, M., 2009. On the volatility and comovement of US financial markets around macroeconomic news announcements. J. Financ. Quant. Anal. 44 (6), 1265–1289.
Brunnermeier, M.K., 2009. Deciphering the liquidity and credit crunch 2007–2008. J. Econ. Perspect. 23 (1), 77–100.
Campbell, J.Y., Thompson, S.B., 2008. Predicting excess stock returns out of sample: can anything beat the historical average? Rev. Financ. Stud. 21 (4), 1509–1531.
Carter, C.K., Kohn, R., 1994. On Gibbs sampling for state space models. Biometrika 541–553.
Carvalho, C.M., West, M., 2007. Dynamic matrix-variate graphical models. Bayesian Analysis 2 (1), 69–98.
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. Springer, pp. 27–34.
Casarin, R., Sartore, D., Tronzano, M., 2018. A Bayesian Markov-switching correlation model for contagion analysis on exchange rate markets. J. Bus. Econ. Stat. 36 (1), 101–114.
Ciccarelli, M., Rebucci, A., 2003. Measuring Contagion with a Bayesian, Time-Varying Coefficient Model. European Central Bank. Working paper No. 263.
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. J. Financ. Econom. 5 (2), 285–320.
Chong, C., Kluppelberg, C., 2018. Contagion in financial systems: a Bayesian network approach. SIAM J. Financ. Math. 9 (1), 28–53.
Cogley, T., Sargent, T.J., 2005. Drifts and volatilities: monetary policies and outcomes in the post WWII US. Rev. Econ. Dynam. 8 (2), 262–302.
Cohen, L., Frazzini, A., 2008. Economic links and predictable returns. J. Financ. 63 (4), 1977–2011.
Connolly, R.A., Stivers, C., Sun, L., 2007. Commonality in the time-variation of stock-stock and stock-bond return comovements. J. Financ. Mark. 10 (2), 192–218.
Corsetti, G., Pericoli, M., Sbracia, M., 2011. Correlation analysis of financial contagion. In: Financial Contagion: the Viral Threat to the Wealth of Nations. Wiley, pp. 11–33.
Dornbusch, R., Park, Y.C., Claessens, S., 2000. Contagion: understanding how it spreads. World Bank Res. Obs. 15 (2), 177–197.
Dungey, M., Fry, R., Gonzalez-Hermosillo, B., Martin, V., 2007. Contagion in international bond markets during the Russian and the LTCM crises. J. Financ. Stab. 2 (1), 1–27.
Eickmeier, S., Lemke, W., Marcellino, M., 2015. Classical time varying factor-augmented vector auto-regressive models—estimation, forecasting and structural analysis. J. R. Stat. Soc. Ser. A 178 (3), 493–533.
Flavin, T.J., Sheenan, L., 2015. The role of US subprime mortgage-backed assets in propagating the crisis: contagion or interdependence? N. Am. J. Econ. Finance 34, 167–186.
Forbes, K.J., Rigobon, R., 2002. No contagion, only interdependence: measuring stock market comovements. J. Financ. 57 (5), 2223–2261.
Frühwirth-Schnatter, S., 2006. Finite Mixture and Markov Switching Models. Springer Science & Business Media.
Gerlach, R., Carter, C., Kohn, R., 2000. Efficient Bayesian inference for dynamic mixture models. J. Am. Stat. Assoc. 95 (451), 819–828.
Giordani, P., Kohn, R., 2008. Efficient Bayesian inference for multiple change-point and mixture innovation models. J. Bus. Econ. Stat. 26 (1), 66–77.
Gonzalo, J., Olmo, J., 2005. Contagion versus Flight to Quality in Financial Markets. Working paper. Universidad Carlos III Madrid.
Gorton, G.B., 2009. Information, Liquidity, and the (Ongoing) Panic of 2007. Working paper No. 14649. National Bureau of Economic Research.
Guidolin, M., Pedio, M., 2017. Identifying and measuring the contagion channels at work in the European financial crises. J. Int. Financ. Mark. Inst. Money 48, 117–134.
Guidolin, M., Tam, Y.M., 2013. A yield spread perspective on the great financial crisis: break-point test evidence. Int. Rev. Financ. Anal. 26, 18–39.
Guo, F., Chen, C.R., Huang, Y.S., 2011. Markets contagion during financial crisis: a regime-switching approach. Int. Rev. Econ. Financ. 20 (1), 95–109.
Harrison, J., West, M., 1997. Bayesian Forecasting and Dynamic Models. Springer, New York.
Hartmann, P., Straetmans, S., de Vries, C., 2004. Asset market linkages in crisis periods. Rev. Econ. Stat. 86 (1), 313–326.
Helwege, J., Turner, C.M., 1999. The slope of the credit yield curve for speculative-grade issuers. J. Financ. 54 (5), 1869–1884.
Hong, Y., Tu, J., Zhou, G., 2006. Asymmetries in stock returns: statistical tests and economic evaluation. Rev. Financ. Stud. 20 (5), 1547–1581.
Jackson, M.O., 2014. Networks in the understanding of economic behaviors. J. Econ. Perspect. 28 (4), 3–22.
Jawadi, F., Louhichi, W., Cheffou, A.I., 2015. Testing and modeling jump contagion across international stock markets: a nonparametric intraday approach. J. Financ. Mark. 26, 64–84.
Kaabia, O., Abid, I., Guesmi, K., 2013. Does Bayesian shrinkage help to better reflect what happened during the subprime crisis? Econ. Modell. 31, 423–432.
Kaminsky, G.L., Reinhart, C.M., Vegh, C.A., 2003. The unholy trinity of financial contagion. J. Econ. Perspect. 17 (4), 51–74.
Kelly, B., Lustig, H., Van Nieuwerburgh, S., 2013. Firm Volatility in Granular Networks. National Bureau of Economic Research working paper No. 19466.
Kim, C.J., Nelson, C.R., 1999. State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications. MIT Press.
King, M.A., Wadhwani, S., 1990. Transmission of volatility between stock markets. Rev. Financ. Stud. 3 (1), 5–33.
Kodres, L.E., Pritsker, M., 2002. A rational expectations model of financial contagion. J. Financ. 57 (2), 769–799.
Kohonen, A., 2014. Transmission of government default risk in the eurozone. J. Int. Money Financ. 47, 71–85.
Koop, G., Korobilis, D., 2010. Bayesian multivariate time series methods for empirical macroeconomics. Foundations and Trends in Econometrics 3 (4), 267–358.
Koop, G., Leon-Gonzalez, R., Strachan, R.W., 2009. On the evolution of the monetary policy transmission mechanism. J. Econ. Dyn. Control 33 (4), 997–1017.
Koop, G., Pesaran, M.H., Potter, S.M., 1996. Impulse response analysis in nonlinear multivariate models. J. Econom. 74 (1), 119–147.
Krishnamurthy, A., Vissing-Jorgensen, A., 2012. The aggregate demand for Treasury debt. J. Political Econ. 120 (2), 233–267.
Li, M., Milne, F., Qiu, J., 2016. Uncertainty in an interconnected financial system, contagion, and market freezes. J. Money Credit Bank. 48 (6), 1135–1168.
Longstaff, F.A., 2010. The subprime credit crisis and contagion in financial markets. J. Financ. Econ. 97 (3), 436–450.
Marron, J., Wand, M., 1992. Exact mean integrated squared error. Ann. Stat. 20 (2), 712–736.
Menzly, L., Ozbas, O., 2010. Market segmentation and cross-predictability of returns. J. Financ. 65 (4), 1555–1580.
Mumtaz, H., Surico, P., 2009. The transmission of international shocks: a factor-augmented VAR approach. J. Money Credit Bank. 41 (1), 71–100.
Pasquariello, P., 2014. Financial market dislocations. Rev. Financ. Stud. 27 (6), 1868–1914.
Pericoli, M., Sbracia, M., 2003. A primer on financialcontagion. J. Econ. Surv. 17 (4), 571–608.
Pesaran, M.H., Pick, A., 2007. Econometric issues in the analysis of contagion. J. Econ. Dyn. Control 31 (4), 1245–1277.
Pesaran, M.H., Timmermann, A., 1995. Predictability of stock returns: robustness and economic significance. J. Financ. 50 (4), 1201–1228.
Primiceri, G.E., 2005. Time varying structural vector autoregressions and monetary policy. Rev. Econ. Stud. 72 (3), 821–852.
Rigobon, R., 2016. Contagion, Spillover and Interdependence. European Central Bank working paper No. 1975.
Sensier, M., Van Dijk, D., 2004. Testing for volatility changes in US macroeconomic time series. Rev. Econ. Stat. 86 (3), 833–839.
Sharifkhani, A., Simutin, M., 2016. Feedback Loops in Industry Trade Networks and the Term Structure of Momentum Profits. Working paper. University of Toronto.
Stock, J.H., Watson, M.W., 1996. Evidence on structural instability in macroeconomic time series relations. J. Bus. Econ. Stat. 14 (1), 11–30.
Tu, J., 2010. Is regime switching in stock returns important in portfolio decisions? Manag. Sci. 56 (7), 1547–1581.
Yang, Z., Zhou, Y., 2016. Quantitative easing and volatility spillovers across countries and asset classes. Manag. Sci. 63 (2), 333–354.
Yang, J., Zhou, Y., 2013. Credit risk spillovers among financial institutions around the global credit crisis: firm-level evidence. Manag. Sci. 59 (10), 2343–2359.