NBER WORKING PAPER SERIES

THE GREAT DEPRESSION AND THE GREAT RECESSION:
A VIEW FROM FINANCIAL MARKETS

Francesco Bianchi

Working Paper 21056
http://www.nber.org/papers/w21056

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2015, Revised December 2018

I am grateful to Chris Sims for his useful suggestions at the early stage of this work. I would like to thank Robert Barro, Markus Brunnermeier, John Campbell, Emmanuel Farhi, Andrew Foerster, Jakub Jurek, Ralph Koijen, Howard Kung, Alisdair McKay, Kristo¤er Nimark, Barbara Rossi, Motohiro Yogo, Adam Zawadowski, and seminar participants at the Annual Asset Pricing Retreat, Princeton University, Duke University, the Board of Governors of the FRS, the EEA-ESEM conference, the Bank of Italy, the University of Pisa, and the CEF Conference for helpful discussions and comments. Zhao Liu provided outstanding research assistance. A previous version of this paper circulated under the title “Rare Events, Financial Crises, and the Cross-Section of Asset Returns.” The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

The Great Depression and the Great Recession: A View from Financial Markets
Francesco Bianchi
NBER Working Paper No. 21056
March 2015, Revised December 2018
JEL No. C32,G01,G12

ABSTRACT

Similarities between the Great Depression and the Great Recession are documented with respect to the behavior of financial markets. A Great Depression regime is identified by using a Markov-switching VAR. The probability of this regime has remained close to zero for many decades, but spiked for a short period during the most recent financial crisis, the Great Recession. The Great Depression regime implies a collapse of the stock market, with small-growth stocks outperforming small-value stocks. A model with financial frictions and uncertainty about policy makers' intervention suggests that policy intervention during the Great Recession might have avoided a second Great Depression. A multi-country analysis shows that the Great Depression and Great Recession were not like any other financial crises.

Francesco Bianchi
Social Sciences Building, 201B
Department of Economics
Duke University
Box 90097
Durham, NC 27708-0097
and CEPR
and also NBER
francesco.bianchi@duke.edu

1 Introduction

The recent financial crisis had pervasive consequences, leading the U.S. economy to its longest and most severe recession since World War II. Arguably, the crisis started with the end of a housing bubble that, in turn, led to the collapse of the subprime market. From there, in the span of a few months, it spread to the whole banking sector and then to the real economy. The decline in real economic activity accelerated in the fall of 2008 as the financial crisis unfolded. U.S. gross domestic product fell by 5% in a year, while the unemployment rate increased from less than 5% to 10%. The large contraction in real activity came with an equally dramatic decline in stock prices, with the S&P 500 index dropping by almost 57% from its October 2007 peak of $1,565$ to a mere 676.5 in March 2009. The possibility of a complete financial meltdown suddenly became a real concern and commentators and policymakers alike feared that the economy could be heading toward a second Great Depression (Krugman, 2009).

This paper is interested in studying to what extent the two events are in fact similar by focusing on the behavior of financial markets. The stock market and real economy are not always in sync. The stock market presents large fluctuations both at low and high frequencies that are not immediately reconcilable with the behavior of the real economy. However, both the Great Depression and the Great Recession showed a strong connection between the stock market and the real economy, as shown in Figure 1. The first two panels report the evolution of the price-earnings ratio and industrial production over the first four years of the Great Depression and the Great Recession. In both cases, the starting points are normalized to 100. The starting dates are August 1928 and April 2008 for the Great Depression and the Great Recession, respectively. As pointed out by Eichengreen and O'rourke (2010), over the first year, the two events looked remarkably similar, with both the stock market and industrial production experiencing rapid declines. However, after these initial drops, both real activity and the stock market recovered fairly quickly during the Great Recession, while the same cannot be said about the Great Depression.

The strong connection between the two series is even more evident if we focus on fluctuations at business cycle frequencies. These are obtained with a bandpass filter and are reported in the third panel of Figure 1. While stock market fluctuations are not always in sync with the real economy, both during the Great Depression and the Great Recession the commovement is remarkably strong. This is made evident by the last panel of the figure that reports the correlation between the two series at business cycle frequencies over a 10-year moving window. Over the sample this correlation can be quite low or even negative, but during the Great Depression and the Great Recession it was remarkably close to 1. This suggests that the study of financial markets can provide valuable information about the similarities and differences between the two events.

In order to formally assess to what extent financial markets during the Great Recession mirrored their behavior during the Great Depression, I first estimate a Markov-switching vector autoregression (MS-VAR) that allows for both changes in the VAR coefficients and in the covari-

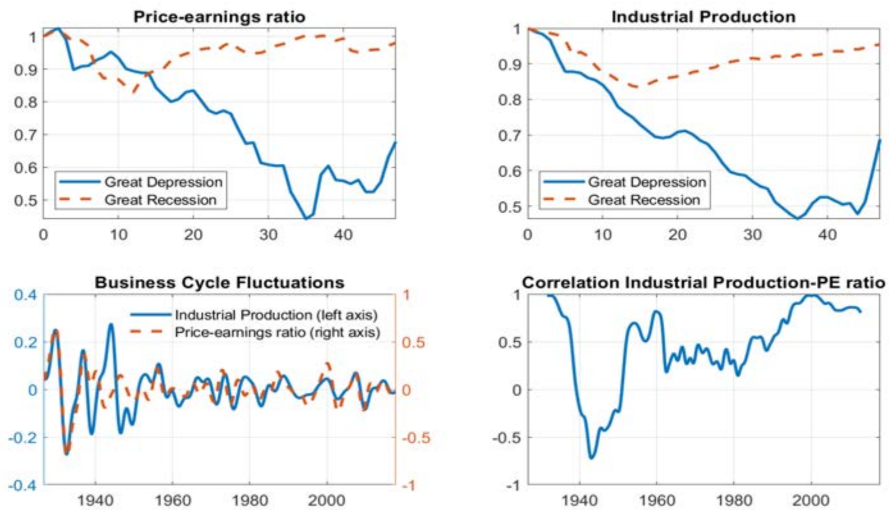

Figure 1: **Stock market and real activity during the Great Depression and the Great Recession.** The first row reports the evolution of the Price-earnings ratio and industrial production during the Great Depression and the Great Recession. The starting points are normalized to 100. The starting dates are August 1928 and April 2008 for the Great Depression and the Great Recession, respectively. The left panel in the second row reports the evolution of the two variables at business cycle frequencies. The right panel in the second row computes the correlation betweent the two series at business cycle frequencies over a 10-year moving window.

ance matrix that characterizes the contemporaneous relations and volatilities of the disturbances. I include four key financial variables: the excess market return, the Term Yield spread, the Price Earnings ratio, and the Value spread. The excess market return captures the performance of the stock market with respect to a risk-free rate. The Term Yield spread measures the slope of the term structure of interest rates that in turn has predictive power for future real activity. The Price Earnings ratio can be considered a measure of market imbalance as it tends to be negatively correlated with future returns. Finally, the Value spread measures the difference between the log book-to-market ratios of small value and small growth stocks. Given that this last variable moves up when small growth stocks perform relatively better, it can be considered a proxy for the behavior of the cross section of asset returns.

A *Great Depression regime* emerges from the estimates. A central feature of this regime is that it implies a large collapse of the stock market with a contemporaneous large increase in the Value spread, suggesting that growth stocks perform relatively better than value stocks during financial crises. As implied by its name, this regime characterized the behavior of the stock market during the Great Depression, when the Price Earnings ratio and the Value spread touched the historical minimum and maximum, respectively. For the remainder of the sample, its probability has been close to zero until the early months of 2009. Therefore, the Great Recession shows a resurgence of this regime, even if for only two months. The probability of the Great Depression regime crossed the threshold of 50% in February 2009 for the first time since November 1948. However, it quickly returned to zero in March, arguably because of government interventions that were effective in preventing a financial meltdown and led to a reversal in the behavior of the stock market and the Value spread.

In order to reinforce this point, I use counterfactual simulations to show that since the starting of the Great Recession in mid-2008 until February 2009, financial markets were on a path consistent with the Great Depression regime: a persistent fall in the stock market paired with a contemporaneous increase in the Value spread. This opposite moving relation between the behavior of the stock market as a whole and the relative performance of growth stocks was absent during another important market decline: The end of the Information Technology (IT) bubble. In that case, the Value spread and the Price Earnings ratios were moving together. This suggests that market declines that are associated with financial crises might be inherently different and that monitoring the relative performance of growth and value portfolios during these events might be useful in understanding where markets are headed.

The similarities between the Great Depression and the Great Recession extend beyond the level dynamics that are implied by the VAR coefficients. Even the innovations present some interesting features. First, both periods were characterized by high volatility. More interestingly, both during the stock market crash that opened the Great Depression and the fall in the stock market that characterized the beginning of the Great Recession, shocks to market returns and the Value spread were negatively correlated. This has an important implication for asset pricing because it implies that during crises, innovations to the relative return of growth stocks move in an opposite direction with respect to stock market returns.

The second part of the paper is devoted to substantiating the idea that policy intervention during the Great Recession might have been key to avoid a second Great Depression. I argue that policy intervention was crucial because it restored the functioning of the financial sector. To make this point, I extend the model by Gertler and Karadi (2011) in two ways. First, I allow for the possibility of a large shock to the parameter controlling the limits to financial intermediaries' leverage ability. This has the effect of generating a drastic and sudden reduction in bank lending, a recession, and a fall in asset values. Second, I allow for uncertainty about the way policymakers will react to this shock. In particular, I allow for uncertainty about whether unconventional monetary policy, broadly defined, will be implemented or not. I show that a policy intervention helps in mitigating the recession and it has an immediate effect on asset values very similar to what presented in Figure 1. This result provides an explanation for why the Great Recession started looking different from the Great Depression once the Troubled Asset Relief Program was introduced.

The importance of the policy response can also be useful to understand the different experiences of developed economies during financial crises. Section 5 looks at financial crises across space and time. From this analysis, a series of interesting results emerge. First, the Great Depression and the Great Recession were unique to the extent that were global phenomena. They both originated in the United States and spread to many developed economies. Thus, their effects on the economies that were affected were substantially more severe when compared to other financial crises. Second, while the US economy and stock market recovered relatively quickly from the 2008 financial crisis, the same cannot be said of many European economies. Arguably the difference is due to the very different policy responses observed in the two economic areas. Third, the pattern of a stock market decline paired with an increase in the value spread during the Great Recession seems common to other modern economies.

Given that the results on the dynamics of the Value Spread point toward the existence of an interesting link between major financial crises and the cross section of asset returns, I devote the last part of the paper to a detailed analysis of the implications of the Great Depression and the Great Recession for the cross section of asset returns. I reconsider the *Bad Beta, Good Beta* Intertemporal CAPM (ICAPM) proposed by Campbell and Vuolteenaho (2004). The model is based on the idea that unexpected excess returns can be decomposed into news about future cash flows and news about future discount rates. The ICAPM predicts that the price of risk for the discount-rate beta should equal the variance of unexpected market returns, while the price of risk for the cash-flow beta should be times greater, where is the investor's coefficient of relative risk aversion.

As a first step, the VAR methodology used to derive the news is extended in order to reflect the possibility of regime changes. The ICAPM is then tested over different subsamples to highlight the importance of the two financial crises. Specifically, I use moving windows of 35 years starting from the late 1920s until the recent crisis. The results provide support for the idea that severe financial crises play an important role in explaining the cross section of asset returns. During the early years of the sample, the ICAPM performs well in explaining the 25 Fama-French portfolios sorted with respect to size and book-to-market ratios. However, as the data window moves away from the Great Depression, the explanatory power of the ICAPM starts to slowly decline. However, as the window approaches the most recent financial crisis, the explanatory power of the ICAPM increases steeply, and the R^2 touches 60%, a value that was last reached at the end of 1978.

In order to highlight why the Great Recession plays such an important role in improving the fit of the ICAPM, I show that the return of medium size growth stocks was visibly lower than the expected return implied by the ICAPM during the 1980s and 1990s. In other words, the return on these stocks was too low in light of a general increase in their risk level as captured by their discount rate and cash-flow betas. Symmetrically, returns on value stocks were quite high with respect to what was predicted by the ICAPM. In both cases, the anomalies were largely corrected during the Great Recession. This result suggests that the relative performance of these two classes of stocks during regular times might be compensated by their behavior during financial crises.

Furthermore, financial crises are also important in shaping agents' expectations. This conclusion can be inferred by comparing the explanatory power of the ICAPM under the benchmark case, in which *fully rational* agents form expectations taking into account the possibility of regime changes, with an alternative scenario in which agents form expectations disregarding the possibility of regime changes. This latter case corresponds to the case of anticipated utility: at each point in time, agents assume that the probabilities of the two regimes will not change in the future. I show that the benchmark case in which agents consider the existence of the Great Depression regime delivers substantially better results.

As a methodological contribution, this paper proposes a simple algorithm to estimate a Markov-switching VAR in reduced form with Bayesian methods. An MS-VAR allows for an analytical characterization of the news along the lines of the VAR approach proposed by Campbell (1991) to implement the present value decomposition of Campbell and Shiller (1988). The formulas presented in the paper can be easily modified to handle other models that make use of a present value decomposition to allow for the possibility of structural breaks. This approach, which formally isolates periods characterized by unusual dynamics, might also prove useful in explaining why the present value decomposition methodology is often sensitive to the sample choice. Furthermore, the Markov-switching extension can easily accommodate temporary non-stationary regimes as long as the system as a whole is stable.

I argued above that the financial crises that coincided with the Great Depression and the Great Recession are different from other financial crises. From this point of view, they can be considered rare events. Thus, this paper is also related to the growing rare disasters literature (Rietz (1988), Barro (2006, 2009), Nakamura et al. (2013), Gabaix (2012), Bollerslev and Todorov (2011), Wachter (2013), Gourio (2012), Bai et al. (2015), Julliard and Ghosh (2012), among others).

The paper is also connected to the vast literature on the cross section of asset returns. Zhang (2005) shows that the value anomaly arises naturally due to costly reversibility and the countercyclical price of risk. Campbell et al. (2013) highlight that the 2007-2009 market fall was not offset by improving stock return forecasts as in the stock market downturn of 2000-2002, while Campbell et al. (2014) extend the approximate closed-form intertemporal capital asset pricing model of Campbell (1993) to allow for stochastic volatility. With respect to their work, I do not impose the restriction that all volatilities have to move in parallel, I allow the covariance structure of the disturbances to vary over time, and I model the possibility of regime changes in the VAR coefficients and, consequently, in the way agents map shocks into the news about future discount rates and future cash flows. To the best of my knowledge, this feature is new in the literature. Given that the primary interest of this paper is to assess the role of the Great Recession and the Great Depression, I do not price volatility when studying the cross section of asset returns. The possibility of merging the two approaches is an interesting path for future research.

Markov-switching models are quite popular in financial econometrics. See Lettau et al. (2008), Ang and Bekaert (2002), Pesaran et al. (2006), Gulen et al. (2011), Gulen et al. (2011), and Bianchi et al. (2016) among others. With respect to these contributions, I use a multivariate model with two separate processes controlling the VAR coefficients and the volatilities, while the literature often utilizes univariate processes in which a single chain controls all parameters of the model.

Allowing for two separate chains is important because volatility changes would otherwise tend to dominate other regime breaks (see Ang and Timmermann (2012) and Sims and Zha (2006)).

The content of this paper can be summarized as follows. In Section 2, I present the MS-VAR used to assess the presence of similarities between the Great Depression and the Great Recession. Section 3 reports the results for the MS-VAR estimates. In Section 4, I present the macro model with financial frictions and policy uncertainty. Section 5 shows that the Great Depression and the Great Recession were not like any other financial crisis. Section 6 presents the implications of the two events for the cross section of asset returns. In Section 7, I conclude.

2 The Model

In this section, I present the MS-VAR use to study the similarities between the Great Depression and the Great Recession.

2.1 A Markov-switching VAR

The Z_t is a $(n \times 1)$ vector of data evolves according to a Markov-switching VAR with one lag:

$$Z_t = c_{\xi_t^\Phi} + A_{\xi_t^\Phi} Z_{t-1} + \Sigma_{\xi_t^\Sigma}^{1/2} \omega_t \tag{1}$$

$$\Phi_{\xi_t^\Phi} = \left[c_{\xi_t^\Phi}, A_{\xi_t^\Phi} \right], \; \omega_t \sim N(0, I) \tag{2}$$

where the unobserved states ξ_t^Σ and ξ_t^Φ can take on a finite number of values, $j^\Phi = 1, \dots, m^\Phi$ and $j^\Sigma = 1, \dots, m^\Sigma$, and follow two independent Markov chains. This represents a convenient way to model heteroskedasticity and to allow for the possibility of changes in the dynamics of the state variables. The probability of moving from one state to another is given by $P[\xi_t^\Phi = i | \xi_{t-1}^\Phi = j] = h_{ij}^\Phi$ and $P[\xi_t^\Sigma = i | \xi_{t-1}^\Sigma = j] = h_{ij}^\Sigma$.

Given $H^\Phi = [h_{ij}^\Phi]$ and $H^\Sigma = [h_{ij}^\Sigma]$ and a prior distribution for the initial state, we can compute the likelihood of the parameters of the model, conditional on the initial observation Z_0. The likelihood can then be combined with a prior probability for the parameters of the model to obtain their posterior probability. A by-product of the likelihood calculation are the filtered probabilities for Markov-switching states: $\pi_{t|t}^\Phi$ and $\pi_{t|t}^\Sigma$, where each element of the two vectors is defined by $\pi_{t|t}^{\Phi,i} = P[\xi_t^\Phi = i | Z^t, \Phi_{\xi_t^\Phi}, \Sigma_{\xi_t^\Sigma}, H^\Phi, H^\Sigma]$ and $\pi_{t|t}^{\Sigma,i} = P[\xi_t^\Sigma = i | Z^t, \Phi_{\xi_t^\Phi}, \Sigma_{\xi_t^\Sigma}, H^\Phi, H^\Sigma]$ for all i at each t where $Z^t = \{Z_s\}_{s=1}^t$. Therefore, the filtered estimates represent the probabilities assigned to the different regimes conditional on the model parameters and the data up to time t. These can be converted by a recursive algorithm to smoothed estimates: $\pi_{t|T}^\Phi$ and $\pi_{t|T}^\Sigma$, where each element of the two vectors is given by $P[\xi_t^\Phi = i | Z^T, \Phi_{\xi_t^\Phi}, \Sigma_{\xi_t^\Sigma}, H^\Phi, H^\Sigma]$ and $P[\xi_t^\Sigma = i | Z^T, \Phi_{\xi_t^\Phi}, \Sigma_{\xi_t^\Sigma}, H^\Phi, H^\Sigma]$. These are probabilities for the different regimes conditional on the model parameters and the whole dataset $Z^T = \{Z_s\}_{s=1}^T$.

2.2 Dataset and Bayesian inference

The vector Z_t contains four state variables: The excess log return on the CRSP value-weighted index (ER_t), the Term Yield spread in percentage points (TY_t), measured as the yield difference between ten-year constant-maturity taxable bonds and short-term taxable notes, the log price earning ratio (PE_t), and the small-stock value-spread (VS_t), the difference in the log book-to-market ratios of small-value and small-growth stocks. The sample spans the period from December 1928 to June 2009.

The construction of the series follows Campbell and Vuolteenaho (2004). The excess market return is computed as the difference between the log return on the Center for Research in Securities Prices (CRSP) value-weighted stock index and the rate on three-month Treasury bills. The Term Yield spread is computed using data available on Global Financial Data by taking the yield difference between 10-year constant-maturity taxable bonds and short-term taxable notes, in percentage points. The Price Earnings ratio (Shiller, 2000) is the log of the ratio between the price of the S&P 500 index and a 10-year moving average of aggregate earnings of companies in the S&P 500 index. In line with the literature, earnings are averaged to avoid spikes in the Price Earnings ratio caused by cyclical fluctuations in earnings. The moving average is lagged by one quarter in order to ensure that all components of the time-t Price Earnings ratio are observable at time t.

The small-stock Value spread is constructed by using the six "elementary" portfolios available on Professor French's website. These elementary portfolios, which are constructed at the end of each June, are the intersections of two portfolios based on size (market equity, ME) and three portfolios formed on the ratio of book equity to market equity (BE/ME). The size breakpoint for year t is the median NYSE market equity at the end of June of year t. The book-to-market ratio for June of year t is the book equity for the last fiscal year end in $t - 1$ divided by ME for December of $t - 1$. The BE/ME breakpoints are the 30th and 70th NYSE percentiles.

At the end of June of year t, the small-stock Value spread is given by the difference between the $\ln(BE/ME)$ of the small high-book-to-market portfolio and the $\ln(BE/ME)$ of the small low-book-to-market portfolio. For months July through May, the small-stock Value spread is updated by adding the cumulative log return from the previous June on the small low-book-to-market portfolio minus the cumulative log return on the small high-book-to-market portfolio to the end-of-June small-stock Value spread. Therefore, an increase in the Value spread reflects the fact that small-growth stocks are outperforming small-value stocks.

The model is estimated with Bayesian methods. Proper priors are put on all the parameters in the model. The priors for all parameters are very loose and identical across the different regimes. This implies that the features of the regimes are not restricted and that differences will arise only because of the data. Appendix A describes the priors in detail. I also impose covariance stationarity by adopting the concept of mean square stability. An MS model is mean square stable if both the first and second moments converge.[1] The posterior is obtained combining the likelihood with the priors. Appendix B describes how to compute the likelihood and the regime probabilities for a given set of parameters. I first search for the posterior mode maximizing the sum of the logarithm of the priors and the log-likelihood. This is an important step because MS models tend to have multiple peaks. I then employ a Gibbs sampling algorithm to draw from the posterior distribution. The algorithm is described in detail in Appendix C. I use 1,000,000 Gibbs sampling iterations of which one every 100 are retained. Convergence is checked using the methods suggested by Geweke (1992) and Raftery and Lewis (1992).

3 The Great Depression and The Great Recession

In what follows I highlight the similarities and differences between the Great Depression and the Great Recession.

3.1 Parameter estimates and regime probabilities

This subsection reports parameter estimates and regime probabilities for the MS-VAR described above. The number of regimes for the VAR coefficients is equal to two, $m^{\Phi} = 2$, while the number of regimes for the covariance matrix is equal to three, $m^{\Sigma} = 3$.[2] Therefore, we have a total of six possible regime combinations. Figure 2 shows the smoothed and filtered probabilities of Regime 1 for the VAR coefficients ($\xi_t^{\Phi} = 1$) at the posterior mode, while Figure 3 reports the smoothed and filtered probabilities for Regime 1 and Regime 3 for the covariance matrices ($\xi_t^{\Sigma} = 1$ and $\xi_t^{\Sigma} = 3$). Table 1 reports posterior mode and 68% error bands for the parameters of the Markov-switching VAR.

I shall start by analyzing the results for the VAR coefficients. The upper panel of Figure 2 contains the filtered and smoothed probabilities of Regime 1 for the VAR coefficients ($\xi_t^{\Phi} = 1$) together with the evolution of the Price Earnings ratio and the Value spread, where the variables have been normalized to fit in the graph. I report both the filtered and smoothed probabilities because they convey different information. We can think about the filtered probability as the probability that would be attached to a particular regime by an agent that was aware of all parameters of the model except for the regime in place at time t. In other words, this is the

[1] Mean square stability holds if and only if all the eigenvalues of the matrix $\Xi \equiv bdiag(A_1 \quad A_1,..., A_m \quad A_m)(H^{\Phi} \quad I_{n^2})$ are inside the unit circle where *bdiag* is a matrix operator that takes a sequence of matrices and construct a block diagonal matrix. Please refer to Costa et al. (2004) and Bianchi (2016) for more details.

[2] I also estimated versions of the model with two and four volatilities regimes. When only two volatilities regimes are considered, I run into the problem that shifts in VAR coefficients are used to compensate for the small number of volatility regimes. This is a problem that has been noted in the literature: When trying to identify "structural" changes in a VAR, it is important to control for stochastic volatility (see, for example, Sims and Zha (2006)). Two regimes do not seem enough to address this issue. When allowing for a fourth regime, I did not find an improvement in fit.

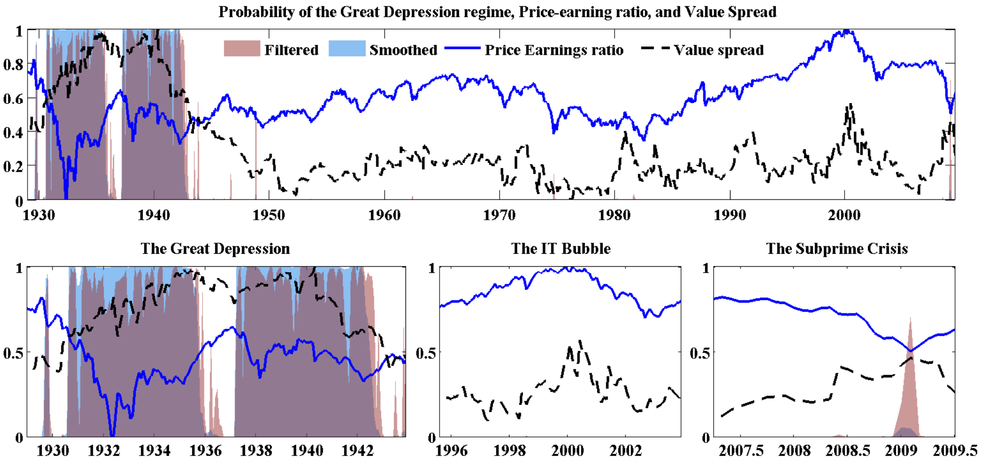

Figure 2: **Regime probabilities for VAR coefficients.** The first panel reports the filtered (red/dark gray area) and smoothed (blue/light gray area) probabilities of the Great Depression regime together with the evolution of the price-earnings ratio and the value spread. The lower three panels zoom on three key events: The Great Depression, the end of the Information Technology bubble, and the Great Recession.

probability that an agent would attach to Regime 1 at time t if she knew the VAR coefficients, the covariance matrices, the transition matrices, and only the data up to time t. Instead, the smoothed probabilities reflect all the information contained in the dataset. This is the probability that an agent would attach to Regime 1 at time t if she knew the VAR coefficients, the covariance matrices, the transition matrices, and the whole dataset up to time T.

In order to facilitate the interpretation of the results, the second row of the figure focuses on three key events: The Great Depression, the IT bubble, and the Great Recession. Regime 1 clearly dominates the first decade, a period characterized by large market crashes and an unusually high level for the Value spread. The behavior of the Value spread and the price earning ratio in the early 1930s is worth noting. The largest stock market crash of U.S. history came with a substantial increase in the Value spread that reached historic heights. In other words, during the most severe recession that the U.S. has ever experienced, growth stocks were outperforming value stocks, and this situation of disequilibrium lasted for more than a decade. The probability of this regime went down only around 1942, when the U.S. started winning WWII. A rational agent who is trying to hedge against risk is likely to find this pattern extremely informative. From here forward I will refer to Regime 1 as the *Great Depression regime*, while I will name Regime 2 the *Regular times regime*.

After the 1930s, the probability of the *Great Depression regime* has generally been close to zero. However, a visible increase in the probability occurred in correspondence with the recent financial crisis. The increase is much larger for the filtered probability than for the smoothed probability. This implies that an agent that had been observing the market in real time would have attached a much larger probability to entering a depression-like regime, while ex-post, with the benefit of the hindsight, the same agent would have concluded that the probability of having observed a manifestation of the Great Depression regime was in fact much smaller. However, even in this latter case, in which the entire dataset is used to infer the smoothed probabilities, we cannot rule out the possibility that during the first two months of 2009, financial markets' behavior was in line with what occurred during the dawn of the Great Depression.

While there are other periods of time during which we observe an increase in the filtered probability of the Great Depression regime, the second month of 2009 was the first time that such a probability crossed 50% since November 1948, a period marked by the rise of the Cold War, the first Israeli-Arab war, and the unexpected presidential election victory of the incumbent President Truman over the Republican candidate, Thomas E. Dewey. Similarly, in the first two months of 2009, the smoothed probability crossed the 5% value for the first time since September 1942, i.e., since World War II. Finally, it is worth emphasizing that these results are even stronger if we were to recursively estimate the model. In this case, the probability assigned to the Great Depression regime would be larger than 80% in February 2009. Appendix F reports results for this alternative approach.

Later, I will investigate more in depth what could explain the increase in the probability of the Great Depression regime at the beginning of 2009. For now, it is enough to point out that the spike in the probability of the Great Depression regime at the beginning of 2009 coincides with a deep decline in the Price Earnings ratio combined with a substantial increase in the Value spread. In other words, the price earning ratio and the Value spread are moving in opposite directions in a way that is very similar to what occurred during the early years of the Great Depression. In this respect, it is quite instructive to compare the Great Recession stock market decline with the end of the IT bubble. In this second case, the Value spread and the Price Earnings ratio were moving in parallel. Recall that the Value spread tends to rise when growth stocks perform relatively better than value stocks. Given that the rise and burst of the IT bubble were mostly driven by IT stocks, it is not surprising that the two variables were moving together. Nevertheless, this evidence suggests that stock market crashes that are associated with financial crises might have very different implications for the relative performance of growth and value stocks. This is why the probability of the Great Depression regime does not increase every time that the Price Earnings ratio falls, but it is more likely to do so if such a fall is associated with a contemporaneous increase in the Value spread.

The two regimes are strongly identified and the parameter estimates present some distinctive features. First of all, the autoregressive component for excess returns is substantially larger under the Great Depression regime ($\xi_t^\Phi = 1$), while the autoregressive components for the Term Yield spread, the Price Earnings ratio, and the Value spread are substantially smaller. A high price earning ratio predicts low stock market returns in both regimes, but the effect is significantly stronger under the Great Depression regime. The Value spread enters the excess return and Price Earnings ratio equations with a positive sign in both regimes, but the coefficients are substantially

$\xi_t^{\Phi} = 1$	ER_t	TY_t	PE_t	VS_t	$constant$
ER_{t+1}	0.1650 (0.0509,0.1822)	□0.0184 (□0.0255,□0.003)	□0.1225 (□0.1232,□0.0937)	0.1275 (0.0894,0.1341)	0.0423 (0.0065,0.0608)
TY_{t+1}	□0.1492 (□0.4170,0.1408)	0.8874 (0.8468,0.9644)	□0.0332 (□0.1030,0.0761)	0.0804 (□0.0958,0.1800)	0.1460 (□0.0512,0.3517)
PE_{t+1}	0.1657 (0.0480,0.1832)	□0.0175 (□0.0247,□0.0007)	0.8732 (0.8720,0.9029)	0.1444 (0.1022,0.1498)	0.0050 (□0.0266,0.0288)
VS_{t+1}	0.0062 (□0.0438,0.0587)	0.0450 (0.0307,0.0490)	0.0563 (0.0365,0.0617)	0.9000 (0.8905,0.9299)	0.0020 (□0.0228,0.0373)
$\xi_t^{\Phi} = 2$	ER_t	TY_t	PE_t	VS_t	$constant$
ER_{t+1}	0.0563 (0.0251,0.0939)	0.0010 (0.0001,0.0029)	□0.0134 (□0.0176,□0.0093)	0.0157 (0.0088,0.0216)	0.0234 (0.0094,0.0380)
TY_{t+1}	0.2591 (0.0351,0.4762)	0.9638 (0.9580,0.9756)	□0.0171 (□0.0416,0.0118)	0.0729 (0.0349,0.1130)	□0.0176 (□0.1263,0.0669)
PE_{t+1}	0.0187 (□0.0125,0.0575)	0.0016 (0.0006,0.0034)	0.9913 (0.9869,0.9952)	0.0193 (0.0121,0.0250)	□0.0019 (□0.0150,0.0138)
VS_{t+1}	□0.0024 (□0.0286,0.0263)	□0.0027 (□0.0033,□0.0010)	□0.0037 (□0.0067,0.0002)	0.9765 (0.9698,0.9812)	0.0465 (0.0331,0.0582)

$\xi_t^{\Sigma} = 1$	u_{ER}	u_{TY}	u_{PE}	u_{VS}
u_{ER}	0.0653 (0.0679,0.0840)	0.0033 (□0.0046,0.0125)	0.0036 (0.0039,0.0061)	0.0015 (0.001,0.0021)
u_{TY}	0.0617 (□0.0657,0.1723)	0.8067 (0.8449,1.0257)	0.0026 (□0.0051,0.0109)	□0.005 (□0.0140,□0.0022)
u_{PE}	0.9189 (0.8864,0.9343)	0.0520 (□0.0795,0.1621)	0.0608 (0.0633,0.0788)	0.0014 (0.0009,0.0020)
u_{VS}	0.4559 (0.256,0.4975)	□0.1212 (□0.2739,□0.0463)	0.457 (0.2596,0.5016)	0.0506 (0.0482,0.0585)
$\xi_t^{\Sigma} = 2$	u_{ER}	u_{TY}	u_{PE}	u_{VS}
u_{ER}	0.0363 (0.0367,0.0389)	□0.0005 (□0.0009,□0.0002)	0.0013 (0.0013,0.0015)	0.0001 (0.0001,0.0002)
u_{TY}	□0.0688 (□0.1056,□0.0259)	0.2133 (0.2233,0.2470)	□0.0005 (□0.0009,□0.0002)	0.0002 (□0.0001,0.0004)
u_{PE}	0.9474 (0.9473,0.955)	□0.0592 (□0.0955,□0.0169)	0.0369 (0.0371,0.0394)	0.0001 (0.0001,0.0002)
u_{VS}	0.1194 (0.1152,0.2006)	0.0285 (□0.0181,0.0604)	0.1001 (0.0944,0.1785)	0.0279 (0.0284,0.0305)
$\xi_t^{\Sigma} = 3$	u_{ER}	u_{TY}	u_{PE}	u_{VS}
u_{ER}	0.1040 (0.1068,0.1270)	□0.0078 (□0.0132,□0.0026)	0.0110 (0.0116,0.0165)	□0.0026 (□0.0051,□0.0016)
u_{TY}	□0.2353 (□0.3151,□0.0651)	0.3183 (0.3083,0.3922)	□0.0085 (□0.0143,□0.0033)	□0.0025 (□0.0076,0.0022)
u_{PE}	0.9611 (0.9521,0.9694)	□0.2411 (□0.3240,□0.0800)	0.1102 (0.1134,0.1348)	□0.0029 (□0.0057,□0.0019)
u_{VS}	□0.2206 (□0.3286,□0.1158)	□0.0684 (□0.1714,0.0509)	□0.2346 (□0.3425,□0.1305)	0.1139 (0.1161,0.1387)

H^{Φ}	$\xi_t^{\Phi} = 1$	$\xi_t^{\Phi} = 2$
$\xi_{t+1}^{\Phi} = 1$	0.9778 (0.9561,0.9831)	0.0050 (0.0038,0.0099)
$\xi_{t+1}^{\Phi} = 2$	0.0222 (0.0169,0.0439)	0.9950 (0.9901,0.9962)

H^{Σ}	$\xi_t^{\Sigma} = 1$	$\xi_t^{\Sigma} = 2$	$\xi_t^{\Sigma} = 3$
$\xi_{t+1}^{\Sigma} = 1$	0.7959 (0.7343,0.8532)	0.0224 (0.0075,0.0232)	0.0526 (0.0353,0.1037)
$\xi_{t+1}^{\Sigma} = 2$	0.1685 (0.0865,0.2030)	0.9243 (0.9182,0.9419)	0.3666 (0.3667,0.5150)
$\xi_{t+1}^{\Sigma} = 3$	0.0356 (0.0288,0.0940)	0.0533 (0.0443,0.0652)	0.5807 (0.4152,0.5636)

Table 1: Parameter estimates. The three sets of tables contain modes and 68% error bands for the posterior distribution of the parameters of the Markov-switching VAR. The first two panels report the estimates for the VAR coefficients. The second set of panels contains the standard deviations of the shocks on the main diagonal, the correlations of the shocks below the main diagonal, and the covariances above the main diagonal. The last tables contain the estimates of the transition matrices.

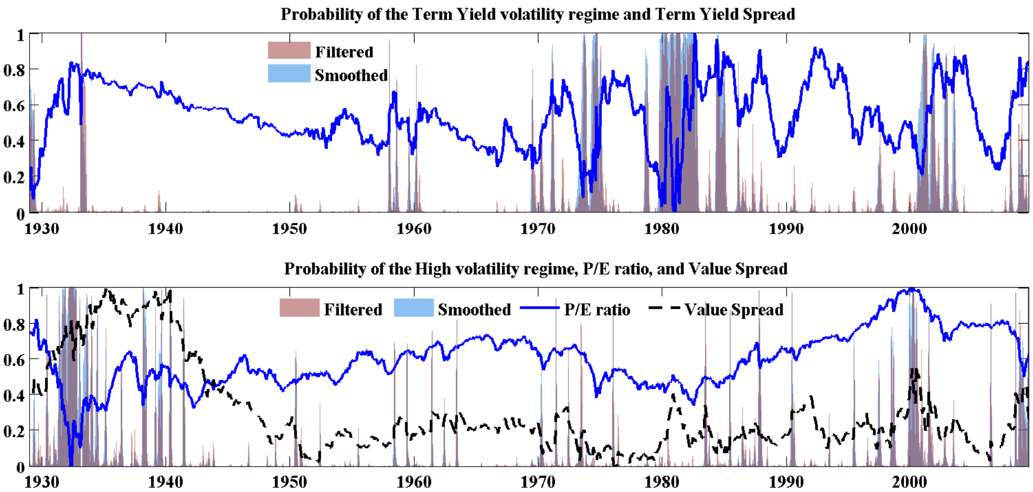

Figure 3: **Probabilities of the volatility regimes.** The figure reports the filtered and smoothed probabilities of Term Yield volatility regime (top panel) and the High volatility regime (lower panel). These two regimes correspond to Regime 1 and Regime 3 for the covariance matrix. The first panel also reports the evolution of the Term Yield Spread, while the second panel contains the Price-earnings ratio and the Value Spread. All variables are rescaled to fit in the 0-1 scale.

larger under the Great Depression regime. Finally, the coefficients of the Term Yield spread and of the Price Earnings ratio in the Value spread equation are positive under the Great Depression regime, while they are smaller and negative in the Regular times regime ($\xi_t^{\Phi} = 2$). Before proceeding, it is worth emphasizing that the dynamic properties of the model do not only depend on the VAR coefficients: Regime changes can also induce strong commovements between the variables of interest. These aspects will be analyzed in the next subsection.

Some interesting patterns emerge from the analysis of the covariance matrix estimates and their corresponding probabilities. Regime 2 ($\xi_t^{\Sigma} = 2$) can be regarded as the *Low volatility regime*, showing the lowest values for the standard deviations of all innovations. Regime 1 ($\xi_t^{\Sigma} = 1$) presents an increase of the magnitude for all shocks, but especially for the innovations to Term Yield spread. Looking at Figure 3, we can see that this regime mostly prevails during the early years of the Volcker chairmanship when the Federal Reserve was targeting reserves with the result of generating high volatility in the FFR and, consequently, the yield spread. I will refer to this regime as the *Term Yield volatility regime*. Regime 3 ($\xi_t^{\Sigma} = 3$) is instead characterized by a more modest increase in the volatility of the Term Yield spread innovations, but a much larger increase in the volatility of the other shocks. Interestingly, Regime 3 prevails for extended periods of time during the 1930s, the 2001 IT bubble burst, and the 2008/9 financial crisis. So it can be considered an *High Uncertainty regime* across several dimensions. The correlation structure of the innovations is also worth noting. Under the High Uncertainty regime, innovations to the Value spread are strongly negatively correlated with innovations to the excess return and Price Earnings equations. This is in sharp contrast with the positive sign that prevails under the other two regimes and implies that small growth stocks tend, in relative terms, to move against the market. Similarly, the correlation of Term Yield innovations with Price Earnings ratio and excess return innovations is strongly negative under the High Uncertainty regime, while under the other two regimes the correlation is slightly negative (Low volatility regime) or centered on zero (Term Yield volatility regime).

Finally, the parameter values for the transition matrix reported at the bottom of Table 1 show that for the VAR coefficients, the Great Depression regime is significantly less persistent and frequent than the Regular times regime, consistent with the idea that the Great Depression was a rare event. As for the transition matrix of the innovation covariance matrix, Regime 2, the Low volatility regime, is the most persistent, followed by the Term Yield volatility regime and the High volatility regime. Their unconditional probabilities are 77.8%, 11.3%, and 10.9%. These estimates imply that the low volatility regime prevails most of the time with relatively frequent but short lasting deviations to the Term Yield volatility regime and the High volatility regime.

3.2 Entering the Great Depression

As mentioned above, regime changes also play a key role in shaping the dynamic properties of the model. In fact, regime changes can be regarded as shocks themselves and can have fairly long lasting consequences. In order to understand the role played by regime changes and at the same time capture the salient features of the Great Depression, Figure 4 reports a simulation in which all Gaussian shocks are set to zero, and regimes follow their most likely path based on the smoothed probabilities at the posterior mode. The initial values coincide with the actual data. The simulated series are reported with a solid blue line, while the red dashed line corresponds to the actual data. The two horizontal lines mark the regime-specific conditional steady states. These are the values to which the variables would converge if a regime were in place forever.

The first aspect that emerges from this simulation is that a change from the Regular times regime to the Great Depression regime determines a sharp drop in the stock market and a contemporaneous increase in the Value spread and the Term Yield spread. The drop in the stock market tends to be very large, and it overshoots with respect to the conditional steady state of the Great Depression regime. Therefore, after a dramatic fall, the stock market partially recovers, while the Value spread and Term Yield spread keep moving toward their corresponding conditional steady states. Notice that the short break in the realization of the Great Depression regime that is identified in the estimates coincides with a partial recovery in the stock market and a contemporaneous fall in the Value spread. However, once the model returns to the Great Depression regime the variables tend to follow a path similar to the one that was prevailing before the break. Overall, during the 1930s the regime sequence plays an important role in tracking the behavior of the three variables, implying that the Great Depression regime captures some salient features of the Great Depression.

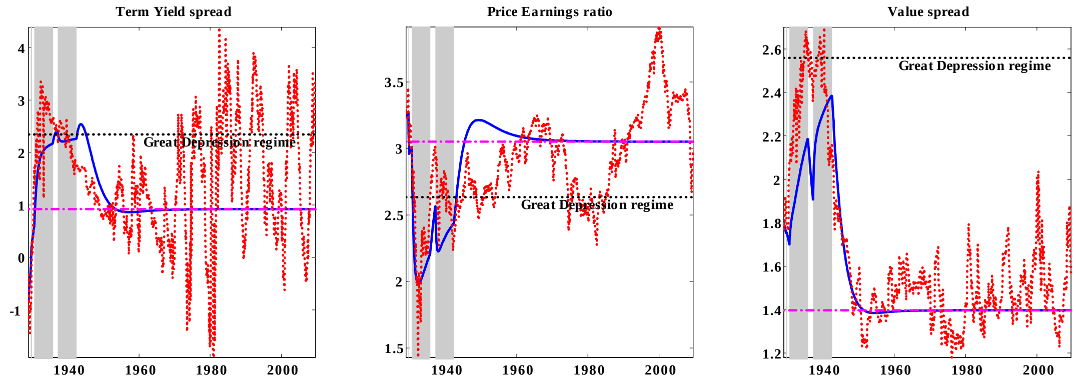

Figure 4: **The Great Depression.** The figure reports a simulation in which all the Gaussian shocks are set to zero, and regimes follow their most likely path based on the smoothed probabilities at the posterior mode. The initial values coincide with the actual data. The simulated series are reported with a solid blue line, while the red dashed line corresponds to the data. The two horizontal lines mark the regime-specific conditional steady states. These are the values to which the variables would converge if a regime were in place forever.

Once the economy returns to the Regular times regime, the model predicts a quick fall in the Value spread and the Term Yield spread. The stock market moves in the opposite direction, showing a steady increase and converging to the higher Regular times conditional steady state. It is also important to emphasize that the conditional steady state for the Great Depression regime is never really reached by the Value spread and the Price Earnings ratio. This is because both in the estimation and in the simulation, the Great Depression regime is not in place long enough to allow for convergence to the conditional steady state.[3]

Figure 5 focuses on the last months of the sample to better understand the similarities and the differences between the Great Depression and the Great Recession. The figure reports two simulations in which all Gaussian shocks have been set to zero starting from February 2009, the month in which the filtered probability of the Great Recession regime spiked. In the first simulation (solid blue line), a counterfactual regime sequence is assumed: starting from February 2009, the Great Depression regime prevails until the end of the sample. In the second simulation, the actual regime sequence is assumed to be in place. The red dotted line corresponds to the data. Notice that until February 2009, the three series coincide. Therefore, the two simulations can be used to understand why the probability of the Great Depression regime increased in the very beginning of 2009, but then quickly fell in March 2009. Furthermore, the simulations shed light on what agents were likely to expect in the moment that the probability of the Great Depression regime spiked.

As already noted above, since the end of 2008 and until February 2009, the stock market experienced a prolonged fall associated with a contemporaneous increase in the Value spread. This behavior is remarkably similar to what is observed in the beginning of the Great Depression.

[3]In Markov-switching models this is a fairly common finding. If the variables converge or not to their conditional steady states depends on the interaction between the persistence of the regime and the persistence of the variables under such a regime.

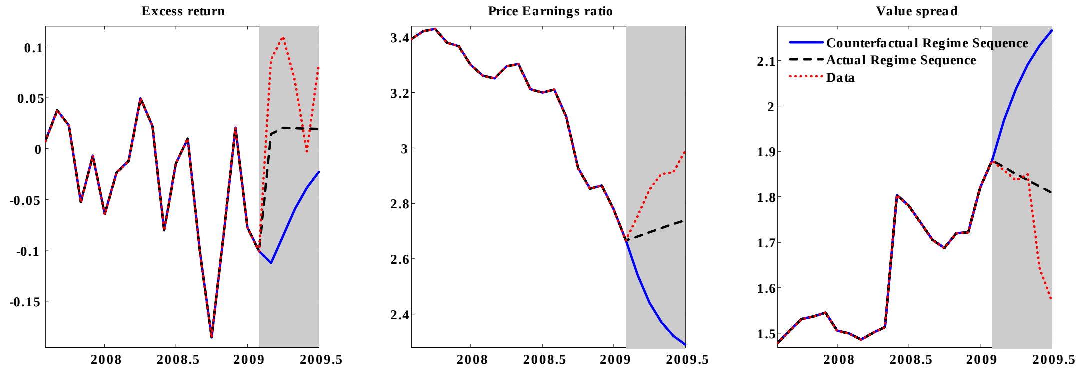

Figure 5: **The Great Recession.** The figure reports two simulations in which all Gaussian shocks have been set to zero starting from March 2009. In the first simulation (solid blue line) a counterfactual regime sequence is assumed: Starting from March 2009 the Great Depression regime prevails until the end of the sample. In the second simulation the actual regime sequence is assumed to be in place. The red dotted line corresponds to the data.

The solid blue line shows what would have happened if starting February 2009 the economy had in fact entered the Great Depression regime: The Price Earnings ratio and the Value spread would have kept moving in exactly the same fashion, while excess returns would have stayed negative. In other words, the counterfactual simulation highlights that until February 2009 financial markets were in fact on a path very similar to what implied by the Great Depression regime. However, in March 2009, these dynamics reverted. Excess stock market returns increased and became positive, the Price Earnings ratio recovered, and the Value spread started declining. The black dashed line shows that the return to the Regular times regime captures these changes, even if in the data the movements were somehow more pronounced. Recall that this is a period of high volatility, so the discrepancy between the "regime only" simulation and the actual data should not be surprising.

In light of these findings, it is then interesting to review the main events that characterized the beginning of the Great Recession. An early flag emerged in June 2007, with the collapse of two hedge funds owned by Bear Stearns. Less than one year later, in March 2008, the Federal Reserve had to intervene in order to prevent the Bear Stearns bankruptcy by assuming $30 billion in liabilities and engineering a sale to JPMorgan Chase. From that moment on, the crisis accelerated with the Treasury Department taking over Fannie Mae and Freddie Mac on September 7, Lehman brothers filing for the largest bankruptcy case in U.S. history one week later (September 15), and the Federal Reserve bailing out AIG. In December 2008, unemployment reached its highest value in 15 years and the Federal Reserve cut the FFR to zero. Over the same period of time, the Price Earnings ratio kept moving down while the Value spread increased.

President Obama took office in January 2009, and Wall Street experienced the worst Inauguration Day drop ever (I am not claiming a causal relation between the two events). At this point, fears that the U.S. might be heading toward a second Great Depression became widespread (Krugman, 2009). On February 10, the secretary of the Treasury Geithner outlined the plan for the expansion of the government bank rescue effort. The plan was received with some skepticism by financial markets, arguably because it was lacking many important details (Solomon, 2009). As a result, the market experienced a fall of almost 5%. A few days later, President Obama signed into law a $787 billion stimulus package that included tax cuts and money for infrastructure, schools, health care, and green energy. Even in this case, some commentators worried that the government intervention might not be enough. In the meantime, the stock market experienced two weeks of declines, reaching its lowest level in 12 years. Notice that it is in February that the probability of the Great Depression regime crossed 50%. However, in March 2009, more encouraging economic data were released and details of the rescue plan were disclosed. Arguably, this had a positive effect of the stock market that turned around. At the same time, the Value spread started moving down and the probability of the Great Depression regime went back to values close to zero.

In summary, the estimates and the counterfactual simulations suggest that during the second half of 2008 and until February 2009, financial markets were on a path consistent with a switch to the Great Depression regime: a falling Price Earnings ratio and an increasing Value spread. This explains the increase in the probability of the Great Depression regime. These patterns came to a stop in March 2009 when the government increased its effort to prevent a financial meltdown and to facilitate an economic recovery. This explains why in the estimates the filtered probability assigned to the Great Depression regime increased significantly at the beginning of 2009, but it quickly went back to zero: The economy did not enter a Great Depression, at least in terms of the behavior of financial markets.

4 Policy Intervention and Asset Valuation

The results presented above suggest that during the Great Recession stock markets behaved in a way consistent with their behavior during the Great Depression until the government outlined a series of policy interventions. One important dimension of these policy interventions was to restore the functioning of the banking sector. In this section, I present a simulation exercise based on the model by Gertler and Karadi (2011) that shows that such policies have the effect of reverting a fall in asset valuation. With respect to the original model, I introduce two ingredients. First, I allow for the possibility of a large shock to the parameter controlling the limits to financial intermediaries' leverage ability. This has the effect of generating a drastic and sudden reduction in bank lending, a recession, and a fall in asset values. Second, I allow for uncertainty about the way policymakers will react to this shock. In particular, I allow for uncertainty about whether unconventional monetary policy will be implemented. I show that unconventional monetary policy, broadly defined, helps in mitigating the recession and it has an immediate effects on asset values.

4.1 The model

The model is based on Gertler and Karadi (2011). I focus on the parts of the model that are different from the original model, while I only provide a brief descriptions of the parts that are directly borrowed from Gertler and Karadi (2011).

Households. The economy is populated by a continuum of households. Within each household there is a fraction f of bankers and a fraction $1 - f$ of workers. The probability of a banker to remain a banker is θ. Bankers and workers engage in perfect consumption sharing within each household. The representative household maximizes:

$$\mathbb{E}_t \sum_{i=0}^{\infty} \beta^i \left[\ln\left(C_{t+i} - hC_{t+i-1}\right) - \chi\left(1+\varphi\right)^{-1} L_{t+i}^{1+\varphi}\right]$$

where C_t and L_t denote consumption and labor, β is the discount factor, h is a parameter controlling habits, and $\varphi > 0$ is the Frisch elasticity of labor supply. Intermediary deposits and government debt are both assumed to be one-period real bonds that pay the gross real return R_t. In equilibrium, the instruments are both riskless and can then be considered perfect substitutes. Thus, this condition is imposed from the beginning. Then the household budget constraint is:

$$C_t = W_t L_t + \Pi_t + T_t + R_t B_t - B_{t+1}.$$

where B_{t+1} is the total amount of short term bonds the household acquires, W_t is the real wage, Π_t corresponds to net payouts to the household from ownership of both non-financial and financial firms, and T_t is a lump sum tax.

Intermediaries. Financial intermediaries use funds obtained from households and their own wealth to lend funds to non-financial firms. The intermediary balance sheet is then given by: $Q_t S_{jt} = N_{jt} + B_{jt+1}$, where N_{jt} is the amount of wealth (net worth) that a banker/intermediary j has at the end of period t, B_{jt+1} are the deposits obtained from households, S_{jt} is the quantity of financial claims on non-financial firms, and Q_t the relative price of each claim. Deposits pay a risk-free rate R_t, while the return $R_{k;t+1}$ of claims on non-financial firms is stochastic return. Thus, net worth N_{jt} follows $N_{j,t+1} = \left(R_{k;t+1} - R_t\right) Q_t S_{j,t} + R_t N_{j,t}$. New bankers receive funds equal to a fraction $\varpi / \left(1 - \theta\right)$ of the assets of exiting bankers.

Intermediaries' participation constraint requires a positive expected discounted spread

$$\mathbb{E}_t \beta^i \left(\varrho_{t+1+i}/\varrho_t\right) N_{j,t+1}\left(R_{k;t+1+i} - R_{t+i}\right) \geq 0 \text{ for } i \geq 0$$

where $\beta^i \left(\varrho_{t+1+i}/\varrho_t\right)$ denotes the household's stochastic discount factor. Intermediaries' terminal net worth is given by

$$V_{j,t} = \mathbb{E}_t \left(1 - \theta\right) \beta \sum_{i=0}^{\infty} \left(\beta\theta\right)^i \left(\varrho_{t+1+i}/\varrho_t\right) N_{j,t+i+1} \tag{3}$$

Thus, the value of being a financial intermediary increases with expected future interest rate spreads, $(R_{k;t+i+1} \square R_{t+i})$, future asset levels $Q_{t+i}S_{j,t+i}$, and the risk-free return on net worth.

As long as the discounted risk adjusted premium is positive, intermediaries will want to expand assets indefinitely. Thus, the model assumes a monitoring problem. Financial intermediaries can divert a time-varying fraction λ_t of its assets back to the household every period, which produces an incentive constraint that requires

$$V_{j,t} \geq \lambda_t Q_t S_{j,t}. \tag{4}$$

The fraction λ_t is given by $\lambda_t = \lambda \exp\left(\widetilde{\lambda}_{\xi_t^\lambda}\right)$, where ξ_t^λ follows a two-state Markov-switching process that jumps between two values: *high* (h) and *low* (l). The probability of moving across the two regimes is controlled by a transition matrix H^λ with diagonal elements equal to p_{hh} and p_{ll}. The values of $\widetilde{\lambda}_h$ and $\widetilde{\lambda}_l$ are such that the unconditional expected value of $\widetilde{\lambda}_{\xi_t^\lambda}$ is 0. This implies that the steady state value of λ_t is λ $(\lambda_{ss} = \lambda)$ and in deviations from steady state we have $\widetilde{\lambda}_t = \widetilde{\lambda}_{\xi_t^\lambda}$. This shock captures the ability of the financial sector to fulfill its tasks of conveying financial resources from households to firms. Most of the time, the economy is in the low state, meaning that only a small fraction of resources can be diverted and the financial sector works properly. During financial crises, λ_t moves to the high state and a large amount of resources can be diverted. In this case, the financial sector ability to transfer resources is jeopardized. As explained below, when the economy is hit by the adverse financial shock, policymakers can react by implementing unconventional monetary policy and mitigate the effects of the shock.

This modeling assumption will create contractions in real activity that originate in the financial sector. In other words, the financial sector does not simply act as a propagation channel, but as an independent source of fluctuations. In this respect, the model is similar to Jermann and Quadrini (2012), while it differs from Gertler and Karadi (2011) and Foerster (2015), where the contraction in real activity is triggered by a shock to the quality of capital. Of course, I could allow for some correlation between this shock and other disturbances affecting the economy. However, it would become harder to disentangle the relative contribution of the different shocks. Importantly, a shock to the capital quality (temporarily) destroys a portion of the capital stock. Instead, a shock to the efficiency of intermediation λ_t leaves the amount of capital available unchanged. Thus, this approach allows me to focus on the functioning of the financial sector, an aspect that is likely to characterize every financial crisis. As shown below, the model goes a long way in generating plausible macro and financial dynamics.

In line with the literature, I assume that the constraint (4) binds at each point in time. Given that all financial intermediaries face this same constraint, total private intermediary assets is given by $Q_t S_{p,t} = \phi_t N_t$, where ϕ_t denotes the leverage ratio. This, in turn, depends negatively on λ_t and positively on the expected discounted marginal gain to intermediaries of expanding assets by a unit, holding net constant, and the expected discounted value of having another unit of net worth holding the amount of claims constant.

The government can also act as a financial intermediary by issuing debt to households and purchasing claims $S_{g,t}$. The government does not face constraints on its balance sheet. However, it might be less efficient than the private sector in providing credit. Such inefficiency is captured by a resource cost of τ for every unit of assets that the central bank owns. The total value Q_tS_t of all assets in the economy is then $Q_tS_t = Q_tS_{p,t} + Q_tS_{g,t}$. The government targets a fraction $\ _t$ of total intermediated assets, so $Q_tS_{g,t} = \ _tQ_tS_t$. Thus, total funds depend on intermediary net worth by $Q_tS_t = \phi_{c,t}N_t$, where the total leverage ratio for the economy $\phi_{c,t} = \phi_t/\left(1 - \ _t\right)$ depends on the amount of intermediation of the government. The policy rule followed by the government to set $\ _t$ is described below.

Non-financial firms. The economy presents three types of non-financial firms: intermediate goods producers, capital producers, and retail firms. Intermediate goods firms produce using capital and labor according to $Y_{m.t} = A_t\left(U_tq_tK_{t-1}\right)^{\alpha} L_t^{1-\alpha}$, where A_t denotes total factor productivity, U_t the capital utilization rate, and q_t is capital quality. Firms purchase capital at price Q_t by issuing claims S_t and hire labor at wage W_t. Capital depreciation rate depends on the utilization rate $\delta\left(U_t\right)$ with elasticity of ζ. The return on capital is given by

$$R_{k;t+1} = q_t\left[P_{m,t}\alpha Y_{m.t}/\left(q_tK_{t-1}\right) + Q_t - \delta\left(U_t\right)\right]/Q_{t-1}$$

which can vary in response to exogenous changes in the capital quality measure q_t. Large fluctuations in the price of claims on capital Q_t generate large swings in the return on capital. These, in turn, determine fluctuations in financial intermediaries' net worth, given that financial intermediaries own the claims.

Competitive capital producers buy capital from intermediate good producing firms, repair depreciated capital, and build new capital. The capital is then sold at price Q_t to the intermediate goods firms. Capital producers face a quadratic adjustment cost on net investment, defined as gross investment less depreciation. The parameter ι controls the inverse of the elasticity of net investment to the capital price. Gross investment I_t equals the total change in capital taking into account depreciation

$$I_t = K_t - \left(1 - \delta\left(U_t\right)\right)q_tK_{t-1}. \tag{5}$$

A continuum of retail firms indexed with $f \in [0,1]$ repackage intermediate output $Y_{m,t}$ into differentiated products $Y_{f,t}$ which they sell at price $P_{f,t}$. Firms face sticky prices a la Calvo with partial indexation to lagged inflation. The probability of reoptimazing the price is $(1 - \ )$. If a firm cannot reoptimize, it sets the price to $P_{f,t} = \Pi_{t-1}^{\mu}P_{f,t-1}$, where $\mu \in [0,1]$ denotes the degree of price indexation to lagged inflation. Steady state net inflation is assumed to be zero. Final output equals a CES aggregate of retail firm goods with elasticity of substitution ε.

Policy rules. The fiscal authority buys a fixed amount of goods $G = gY_{ss}$, where Y_{ss} is steady state output. Furthermore, if the government engages in unconventional monetary policy, it has to pay a cost equal to a fraction τ of the value of its assets. These expanses are financed with lump-sum taxes T_t and the return on previously held assets. Consequently, the government's budget constraint requires

$$G + \tau \quad {}_t Q_t K_t = T_t + \left(R_{k,t} - R_{t-1}\right) B_{g,t-1}$$

Conventional monetary policy sets the nominal interest rate r_t according to a Taylor rule:

$$r_t / r_{ss} = \Pi_t^{\kappa_\pi} \left(Y_t / Y_t^*\right)^{\kappa_y} \exp\left(\sigma_r \varepsilon_{r,t}\right)$$

where r_{ss} denotes the steady state nominal rate, κ_π and κ_y control the responses to inflation and to deviations of output from its flexible-price counterpart Y_t^*, and $\varepsilon_{r,t}$ is a conventional monetary policy shock. The steady-state and target level of inflation is $\Pi_{ss} = 1$.

The government can also conduct unconventional monetary policy. Government asset holding ${}_t$ are controlled by the following rule

$${}_t = \upsilon_{\xi_t^p} \left(\left(\mathbb{E}_t R_{k,t+1} - R_t\right) - \left(R_{k,ss} - R_{ss}\right)\right) + \rho \quad {}_{,\xi_t^p} \quad {}_{t-1}$$

where the response to the expected interest rate spread $\upsilon_{\xi_t^p}$ and the autoregressive term $\rho \quad {}_{,\xi_t^p}$ change according to a two-regime Markov process controlled by ξ_t^p. During a financial crisis spreads increase since the decline in financial intermediaries' net worth limits their ability to take advantage of the spread in returns by acquiring capital claims. By increasing the amount of asset purchases in response to the increase in spreads, the government helps in sustaining credit to the private sector, increasing the price of claims, and, as a result, restoring intermediaries' net worth.

I assume that there are two policy regimes: Conventional and unconventional. Under the conventional monetary policy regime, the government only conducts conventional monetary policy. Thus, $\upsilon_{\xi_t^p} = \upsilon_c = 0$ and the government does not react to the spread. Under the unconventional monetary policy regime, the government *also* conducts unconventional monetary policy. In this case, $\upsilon_{\xi_t} = \upsilon_u > 0$ and the government does react to the spread. I assume that when the financial sector works properly ($\widetilde{\lambda}_{\xi_t^\lambda} = \widetilde{\lambda}_l$), the government only conducts conventional monetary policy. When instead a financial crisis occurs ($\widetilde{\lambda}_{\xi_t^\lambda} = \widetilde{\lambda}_h$), the government *can* intervene to help restoring intermediation of funds from households to firms by moving to the unconventional monetary policy regime. Note that unconventional monetary policy does not need to be exclusively conducted by the central bank.

Importantly, government intervention during a financial crisis is not automatic. When the adverse financial shock $\widetilde{\lambda}_h$ hits, agents cannot be sure that the government will intervene. I model this idea by assuming that in case of a crisis, there is a probability p_u that the government *immediately* implements unconventional monetary policy. If this does not occur, government intervention can still occur with a delay. Specifically, conditional on being in the crisis regime, in every period there is a probability $1 - p_{cc}$ of moving to the unconventional monetary policy regime, where p_{cc} is the persistence of the conventional monetary policy regime conditional on being in the financial crisis regime. Finally, I assume that if the government moves to the unconventional monetary policy regime, it will keep implementing unconventional monetary policy until the financial crisis is over. This boils down to assuming that the persistence of the unconventional monetary policy regime conditional on being in the crisis regime is one ($p_{uu} = 1$).

Summarizing, the following transition matrix characterizes the joint evolution of the shock to the financial sector and policymakers' behavior:

$$H = \left[\begin{array}{c|c} p_{ll} & (1 - p_{hh}) \\ \hline (1 - p_{ll}) \begin{bmatrix} (1 - p_u) \\ p_u \end{bmatrix} & p_{hh} \begin{bmatrix} p_{cc} & 1 - p_{uu} \\ 1 - p_{cc} & p_{uu} \end{bmatrix} \end{array} \right].$$

The combined Markov chain $\xi_t \equiv \left\{ \xi_t^{\lambda}, \xi_t^{p} \right\}$ can assume three values: $\{l, c\}$, $\{h, c\}$, and $\{h, u\}$.

4.2 Solution and Parameterization

The model is linearized around the unique deterministic steady state and solved with the solution method of Farmer et al. (2009). The model solution reflects the fact that agents form expectations taking into account the possibility of a financial crisis and the associated uncertainty about changes in policymakers' behavior. Thus, this modelling framework presents two important features. First, it breaks the orthogonality between shocks and policy that is typically assumed in DSGE models. Here policy can change in response to a particular shock. Second, it captures rational agents' uncertainty about the response of policymakers to the financial crisis. The solution can be characterized as a MS-VAR:

$$S_t = c(\xi_t, \vartheta, H) + T(\xi_t, \vartheta, H) S_{t-1} + R(\xi_t, \vartheta, H) \varepsilon_t \tag{6}$$

where ϑ and S_t are vectors that contain the structural parameters and all the variables of the model, respectively. The law of motion of the model depends on the structural parameters (ϑ), the regime in place (ξ_t), and the probability of moving across regimes (H). This notation highlights that agents' beliefs about future regime changes matter for the law of motion governing the economy.

The parameters used for the simulation are taken from Gertler and Karadi (2011) whenever possible and presented in Table 2. For the transition matrix, I assume that the regular times regime has larger persistence than the financial crisis regime: $p_{ll} = .995$ and $p_{hh} = .95$, respectively. These values imply an average duration of 50 and 5 years, respectively. The probability of policymakers immediately activating unconventional monetary policy in response to a financial crisis is set to 10%. Conditional on being in a financial crisis, in every period there is a 10% probability of

$Parameter$	$Value$	$Parameter$	$Value$	$Parameter$	$Value$	$Parameter$	$Value$
υ_u	20	$\widetilde{\lambda}_l$	□.02	ϖ	0.002	η_i	1.728
p_{ll}	.995	$\widetilde{\lambda}_h$	.20	θ	0.972	α	0.330
p_{hh}	.95	λ	0.381	β	0.990	$\delta(U_{ss})$	0.025
p_u	.1	κ_π	1.5	h	0.815	ζ	7.200
p_{cc}	.9	κ_y	.125	χ	3.409	ε	4.167
p_{uu}	1	ρ_r	.8	φ	0.276		0.779
				g	.2	μ	0.241

Table 2: Parameterization of the microfounded DSGE model.

moving from the conventional monetary policy to the unconventional monetary policy regime (1 □ $p_{cc} = .1$).

Overall, the transition matrix implies that, unconditionally, the economy is in a financial crises with probability around 9% and that the probability of unconventional monetary policy during a crisis is around 50%. The implied frequency and duration of financial crises is roughly in line with what observed in the data, once we take into account that in the model the duration of a financial crisis is meant to capture not just the initial stages, but also the subsequent period of slow recovery and fragile financial system. The probability of unconventional monetary policy during a crisis is in line with the fact that unconventional monetary policy was implemented during the Great Recession, while it was not during the Great Depression.

The response of the government to the spread under the unconventional monetary policy regime υ_u is set to 20, the intermediate value of the ones considered by Gertler and Karadi (2011). As shown below, this value implies that government intervention largely reduces the spread, without completely closing it. Finally, the size of the financial shock across the two regimes is chosen in a way to imply an expected value equal to zero in log-deviations from steady state: $\widetilde{\lambda}_l =$ □.02 and $\widetilde{\lambda}_h = .2$.

Before proceeding, it is worth mentioning that the results presented below are robust to increasing the difference in the persistence of the two regimes, changing the probability of a policy intervention, allowing for the possibility of reversal in the unconventional monetary policy regime. The key ingredient that allows the model to mimic the differences between the Great Depression and the Great Recession is the fact that unconventional monetary policy is not necessarily going to be implemented during a financial crisis.

4.3 Simulation of a financial crisis

This section shows how the parsimonious extension of Gertler and Karadi (2011) presented above can go a long way in understanding the differences between the Great Depression and the Great Recession. Figure 6 presents the response of the economy to a financial crisis with and without policy intervention. All variables are expressed as percentage deviations from steady state. The

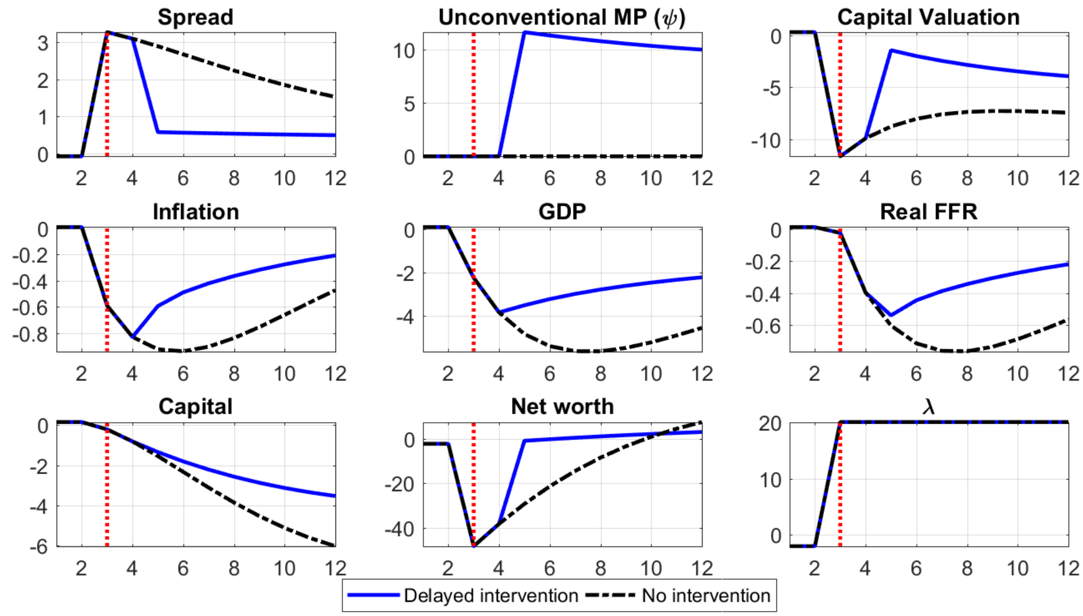

Figure 6: **Crisis and Policy Intervention.** The figure presents the response of the economy to a financial crisis with and without policy intervention. The shock hits the economy in the third period. The solid blue line presents the case in which after two periods the government responds to the crisis by implementing unconventional monetary policy. The black dashed line presents the case in which the government does not intervene.

economy is hit by a financial crisis in period 3. As explained above, this is modelled as an increase in the parameter controlling the amount of funds that bankers can divert ($\widetilde{\lambda}_t = \widetilde{\lambda}_h$). The path for this variable is reported in the lower right corner as the percentage deviation from the steady state. Two cases are then considered with respect to the policy response. In the first case, solid blue line, policymakers react to the shock by implementing unconventional monetary policy starting from period 5. Note that the intervention is delayed with respect to the beginning of the crisis in a way to capture what arguably happened during the Great Recession. In the second case, black dashed line, policymakers do not implement unconventional monetary policy. This second case is meant to capture what happened during the Great Depression, when the Federal Reserve did not implement unconventional monetary policy (see Bernanke (1983)).

Let's consider first the case of no policy intervention (dashed line). In what follows, we use the total capital valuation, $Q_t K_t$, as a proxy for the stock market. The parameter λ_t captures the efficiency of the credit market. If λ_t increases, it becomes harder to recover funds in case bankers try to divert them. As a result, the value of the claims on capital falls precipitously and the credit constraint for financial intermediaries becomes tighter. A process of capital decumulation leads to a progressive reduction in GDP. As the capital stock declines, the overall value of claims that bankers could divert also declines, leading to a progressive reduction in spreads. However, this process is slow and comes with a substantial drop in real activity. This also has a depressing effect on intermediaries' net worth leading to a further reduction in lending. Note that the decline in capital valuation is immediate, while the response of the economy builds-up over time. Finally, like in any new-Keynesian model, nominal rigidities amplify the effect of the shock. The firms that can adjust, lower the price and inflation falls. However, the firms that cannot reoptimize simply set the price following the partial indexation scheme. Thus, the fall in real activity is larger than what it would be under flexible prices.

Government intervention has the effect of reversing these dynamics (blue line). By purchasing assets, the government increases the amount of credit available to the non-financial sector. At the same time, this policy has the effect of increasing the value on capital claims. This helps in increasing banks' net worth with an additional beneficial effect on the amount of credit available. While the effect on the macroeconomy builds-up over time, the effect on asset valuation is immediate. Unconventional monetary policy determines a rapid increase in asset valuations that stabilizes on a higher value, even if the macroeconomy takes some time to recover. Of course, the policy does not completely resolves the problems of the financial sector, as implied by the fact that the spread is not completely reabsorbed. However, the shift in the paths of the macro and financial variables is quite drastic.

These dynamics are remarkably similar to what presented in Figure 1. When policy intervention does not occur, asset valuation experiences a sudden drop and a very slow recovery, while real activity keeps falling for a long time. This pattern seems to characterize the Great Depression. When policy intervention occurs, both the real economy and asset valuation recover, but the latter experiences a substantially faster recovery.

5 Not all Financial Crises are Created Equal

Because of data availability, this paper focuses on the behavior of financial markets for the US economy. However, it is interesting to ask whether the key stylized facts that emerge for the United States can be recovered for other countries and for other financial crises. Figure 7 and Figure 8 present the evolution of industrial production and the stock market for a series of countries during different financial crises. Financial crises dates are identified based on the dataset constructed by Jorda et al. (2016). In line with the previous studies, the authors define systemic financial crises as events during which a country's banking sector experiences bank runs, sharp increases in default rates accompanied by large losses of capital that result in public intervention, bankruptcy, or the forced merger of major financial institutions. The series for the stock market come from the same dataset, while the series for industrial production are obtained from the Global Financial Data website. For each country, the dates for the different financial crises are reported in the legend. In all figures, the solid lines are used to denote the Great Depression and the dashed lines are used to denote the Great Recession. Note that not all countries experienced a financial crisis during these two events and that the dates differ across countries.

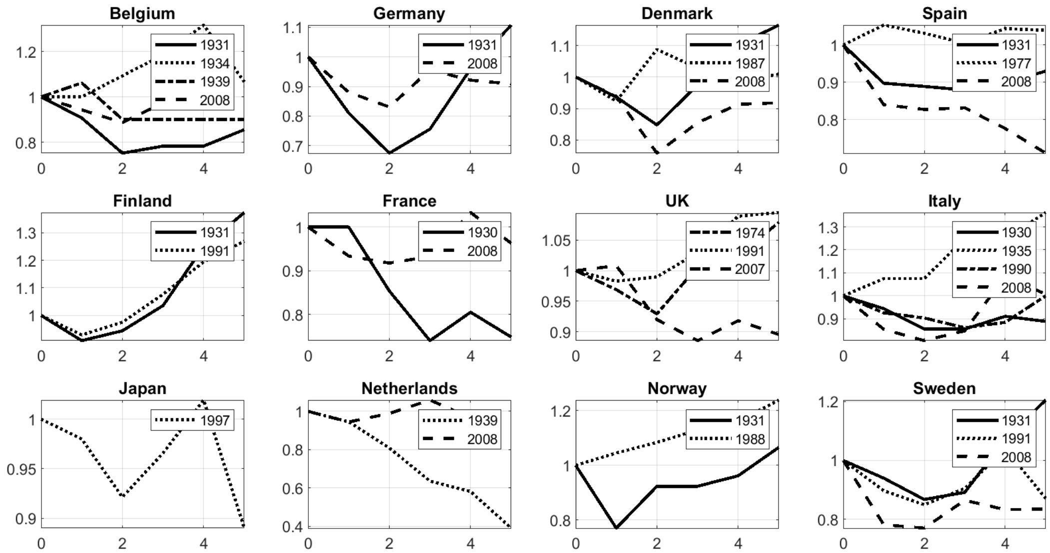

Figure 7: **Real activity and financial crises.** The Figure reports the evolution of industrial production following a financial crisis for a panel of countries. The years of the financial crises are reported in the legends. The solid and dashed lines always correspond to financial crises that occurred during the Great Depression or Great Recession (if applicable), respectively. In all cases, the value of industrial production is mormalized to 1 for the year before the crisis. Thus, in each panel the crisis occurs at time 1.

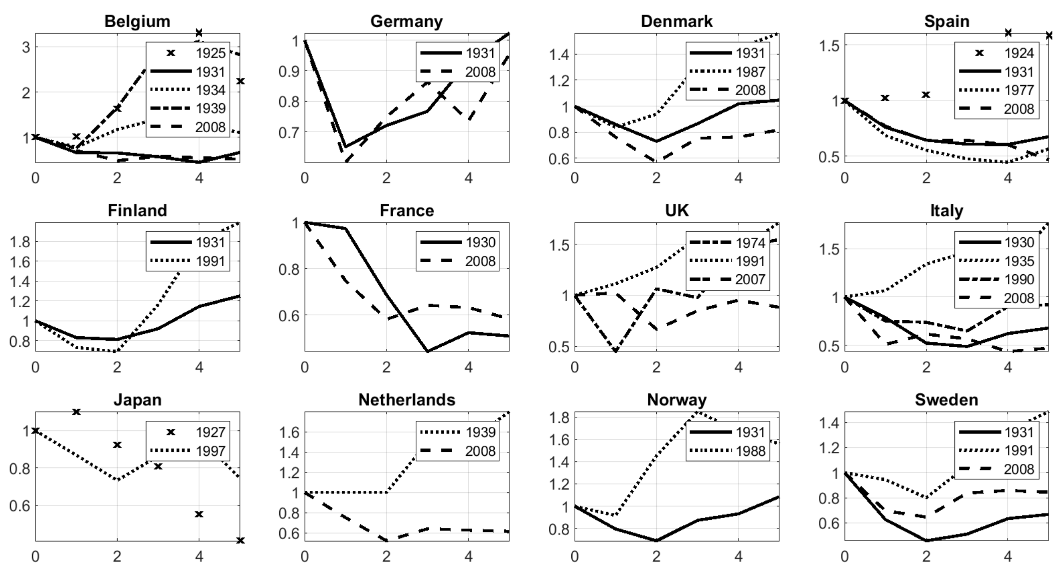

Figure 8: **Stock market and financial crises.** The figure reports the evolution of the stock market following a financial crisis for a panel of countries. The years of the financial crises are reported in the legends. The solid and dashed lines always correspond to financial crises that occurred during the Great Depression or Great Recession (if applicable), respectively. In all cases, the value of industrial production is mormalized to 1 for the year before the crisis. Thus, in each panel the crisis occurs at time 1.

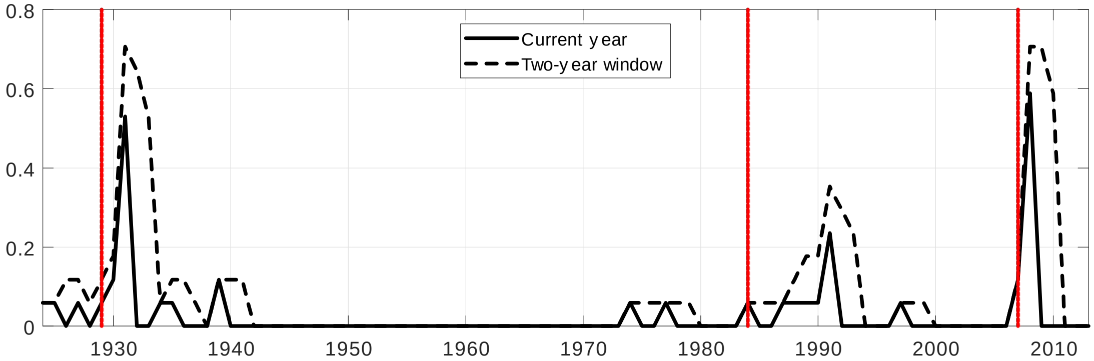

Figure 9: **Fraction of countries in a financial crisis.** The figure reports the fraction of countries that are in a financial crisis in a given year (solid line) or that have been in a crisis in the current or past two years (dashed line) based on the dataset constructed by Jorda, Schularick, and Taylor (2016).

The Great Depression and the Great Recession are always associated with a contraction of industrial production and a decline for the stock market. Instead, other financial crises appear to have less dramatic effects and in many cases we do not observe meaningful declines in industrial production and the stock market. Furthermore, I argued above that in the United States the recovery of both the stock market and industrial production in the aftermath of the Great Recession was quite rapid when compared with the Great Depression. However, the same cannot be said for other large economies. The contractions in industrial production and the decline in the stock market for Germany, Spain, France, United Kingdom, and Italy were large and prolonged and in some cases even larger than during the Great Depression. Such difference in outcomes can be understood in light of the fact that the financial crisis triggered a sovereign debt crisis in Europe. Furthermore, the policy responses in the Euro zone were arguably quite different than in the United States.

Figure 9 provides a possible explanation for why the Great Depression and the Great Recession appear to be much more consequential than other financial crises. The solid line reports the fraction of countries that are in a financial crisis in a given year among the ones that populate the dataset constructed by Jorda et al. (2016). The dashed line computes the fraction of countries that have experienced a financial crisis in the current or previous two years. The vertical lines mark the dates of financial crises that occurred in the United States: 1929 (Great Depression), 1984 (Savings and Loan crisis), and 2007 (Great Recession). Both the Great Depression and the Great Recession were global phenomena, while other financial crises only involved a few countries. Furthermore, both of them originated in the United States and then spread to the rest of the world.

Figure 10 reports the Price-earnings ratio and Value Spread for the United Kingdom, Italy, Spain, Germany, and Japan starting from the early 1990s. Appendix G describes how the data have been constructed. The last panel reports the correlation between the two variables using a 10-year moving window. The dashed vertical line corresponds to the Great Recession (2008), while the dotted vertical line in the panel for Japan corresponds to the 1997 Japanese financial

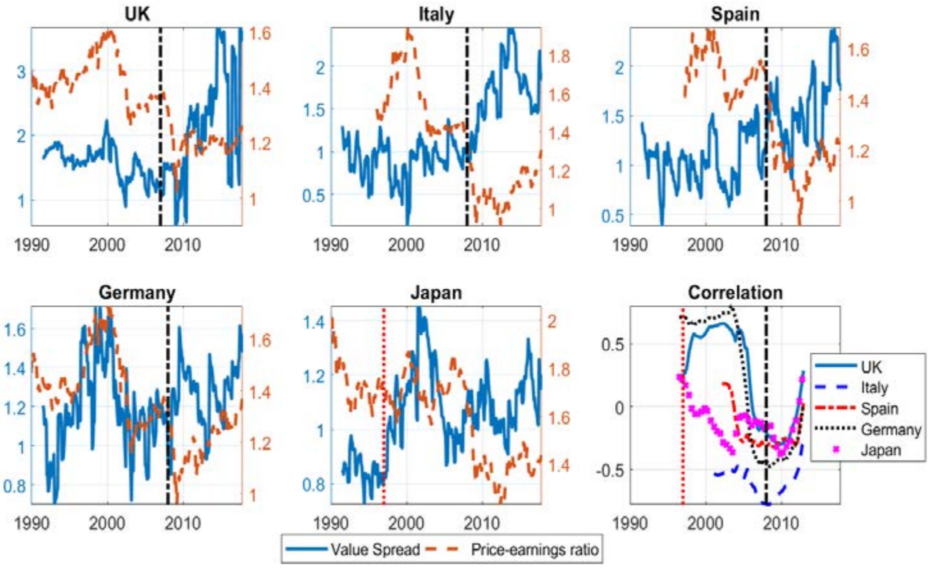

Figure 10: **Price-earnings ratio and Value Spread around the world.** The figure reports the Price-earnings ratio and Value Spread for the United Kingdom, Italy, Spain, Germany, and Japan starting from the early 1990s. The last panel reports the correlation between the two variables uisng a moving window of ten years.

crisis. The Great Recession coincided with a decline in the Price-earnings ratio and an increase in the value spread, in a way very similar to what documented for the United States. In Japan, the negative commovement between the two variables started earlier, following the 1997 financial crisis. Italy presents a negative correlation over the whole sample. With respect to this, it is important to notice that Italy faced a financial crisis in 1990, right before the starting date of the sample. Nevertheless, for all countries that experienced a financial crisis in 2008, the correlation reaches a minimum in correspondence of such event, providing corroborating evidence that financial crises affect both the stock market as a whole, but also the cross section of asset returns. Finally, it is interesting to note that for all countries the value spread remains high, while the Price-earnings ratio did not fully recover, except perhaps for Germany. This result is intriguing because lines up with what presented above for industrial production and stock values: While the US economy and stock market recovered quickly, the same is not true for these other countries.

6 The Cross Section of Asset Returns

The results shown so far have highlighted a series of properties that are quite informative regarding the similarities between the Great Depression and the Great Recession. Two aspects seem particularly relevant. First, during the Great Depression and at the beginning of the Great Recession, the Price Earnings ratio and the Value spread were moving in opposite directions. Second, innovations to the Price Earnings ratio and the Value spread were often negatively correlated during these events. Both results suggest that financial crises imply important changes in the behavior of the cross section of asset returns, with small growth stocks performing relatively better. In order to further explore this idea, I make use of Campbell and Vuolteenaho's ICAPM. Consistent with the Markov-switching model described above, it is important to model the possibility of regime changes when describing agents' expectations formation mechanism. In order to keep the paper self-contained, I will briefly present the ICAPM proposed by Campbell and Vuolteenaho (2004), and then I will explain how to extend their approach to allow for regime changes.

6.1 ICAPM

Fama and French (1992, 1993) show that the CAPM fails to describe average realized stock returns since the early 1960s, when a value-weighted equity index is used as a proxy for the market portfolio. This failure is most apparent for the price of small stocks and value stocks. To solve the small-value puzzle, Campbell and Vuolteenaho (2004) start from the premise that an unexpected change in excess returns can be determined by news about future cash flows or by a change in the discount rate that investors apply to these cash flows. While a fall in expected cash flows is simply bad news, an increase in discount rates implies at least an improvement in future investment opportunities. Therefore, the single CAPM beta can be decomposed into two sub-betas: one reflecting the covariance with news about future cash flows (bad beta), the other linked to news about discount rates (good beta). The previous argument suggests that given two assets with the same CAPM beta, the one with the highest cash-flow beta should have a larger return.

Using the loglinear approximation for returns introduced by Campbell and Shiller (1988), unexpected excess returns can be approximated by:

$$r_{t+1} - \mathbb{E}_t r_{t+1} = N_{CF,t+1} - N_{DR,t+1} = (\mathbb{E}_{t+1} - \mathbb{E}_t) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} - (\mathbb{E}_{t+1} - \mathbb{E}_t) \sum_{j=1}^{\infty} \rho^j r_{t+1+j} \quad (7)$$

where r_{t+1} is a log stock market return, d_{t+1} is the log dividend paid by the stock, Δ denotes a one period change, $\mathbb{E}_t$ denotes a rational expectation formed at time t, and ρ is the discount coefficient that is set to 0.95 per annum. $N_{CF,t+1}$ and $N_{DR,t+1}$ represent news about the future market cash flows and news about the future market discount returns, respectively.

The VAR methodology introduced by Campbell (1991) provides an estimate for the terms $\mathbb{E}_t r_{t+1}$ and $N_{DR,t+1}$. Then $N_{CF,t+1}$ is derived from (7) as a residual. Specifically, consider a VAR in companion form:

$$Z_{t+1} = c + AZ_t + u_{t+1} \quad (8)$$

where Z_t is a vector of state variables with the excess return ordered first. The two types of news can be obtained according to the following transformation of the residuals:

$$r_{t+1} - \mathbb{E}_t r_{t+1} = e_1' u_{t+1},\ \ N_{CF,t+1} = (e_1' + e_1'\lambda)\, u_{t+1},\ \ N_{DR,t+1} = e_1'\lambda u_{t+1} \quad (9)$$

where $\lambda = \rho A (I - \rho A)^{-1}$ and $e_1' = [1, 0, ..., 0]'$. Then, the betas can be computed for a set of portfolios according to the following formulas:

$$\widehat{\beta}_{i,CF} = \frac{\widehat{cov}(r_{i,t}, N_{CF,t})}{\widehat{var}(N_{CF,t} - N_{DR,t})} \text{ and } \widehat{\beta}_{i,DR} = \frac{\widehat{cov}(r_{i,t}, -N_{DR,t})}{\widehat{var}(N_{CF,t} - N_{DR,t})} \tag{10}$$

where $r_{i,t}$ is the return of the *i-th* portfolio. Notice that the denominator is simply the sample variance of the unexpected excess returns, i.e., of the residuals of the VAR $r_{t+1} - \mathbb{E}_t r_{t+1}$. The market beta is obtained by summing the two betas.

Campbell (1993) derives an approximate discrete-time version of Merton's (1973) ICAPM. The pricing implications of the model are based on the first-order condition of an investor with Epstein and Zin (1989) preferences who holds a portfolio of tradable assets that contains all of her wealth. Campbell assumes that this portfolio is observable in order to derive testable asset-pricing implications from the first-order condition. Under appropriate assumptions about the parameters of the model, it can be shown that the price of risk for the discount-rate beta should equal the variance of the market return, while the price of risk for the cash-flow beta should be times greater, where is the investor's coefficient of relative risk aversion.

Three models are examined: the static CAPM, the ICAPM, and an unrestricted factor model based on the two betas. Consider the cross-sectional regression

$$\overline{R}_i = g_0 + g_1 \widehat{\beta}_{i,CF} + g_2 \widehat{\beta}_{i,DR}$$

where $\overline{R}_i$ is the time-series mean for the excess return of asset i. The CAPM model imposes the coefficient restriction $g_1 = g_2$, given that the single market beta is obtained summing the two betas: $\widehat{\beta}_{i,M} = \widehat{\beta}_{i,CF} + \widehat{\beta}_{i,DR}$. According to the ICAPM, the premia should be: $g_1 = \sigma_M^2$ and $g_2 = \sigma_M^2$, where σ_M^2 is the variance of the unexpected excess returns. Therefore, the ICAPM restricts the coefficient of the discount-rate beta, and it returns an estimate of the coefficient of relative risk aversion .[4] In the factor model the coefficients are not restricted. The model can be interpreted as a generalization of the ICAPM that allows the rational investor's portfolio to include Treasury bills as well as equities.

[4] The asset pricing formulas of Campbell (1993) that represent the basis for Campbell and Vuolteenaho (2004) are derived assuming homoskedasticity. However, when modeling parameter instability it is important to allow for heteroskedasticity to avoid spurious changes in the VAR coefficients. This is why the MS-VAR was estimated allowing for heteroskedasticity. Given that the focus here is on the changes in dynamics implied by the Great Depression regime, I regard the idea of extending the analysis to price volatility in the spirit of Campbell et al. (2014) as an interesting direction for future research, but beyond the scope of this paper. Furthermore, even in Campbell and Vuolteenaho (2004) there is not an immediate mapping between the volatility of the VAR innovations (based on estimates obtained over the entire sample) and the variance of the market returns used to price the assets (computed over two distinct subsamples).

6.2 News in a Markov-switching framework

Suppose agents' expectations can be modelled based on the MS-VAR described by (1) and (2). In order to derive the news, we need to be able to model the revision in expectations implied by the MS-VAR residuals taking into account the possibility of regime changes. Define the conditional expectation $\mathbb{E}_0(Z_t) = \mathbb{E}(Z_t|\mathbb{I}_0)$ with $\mathbb{I}_0$ being the information set available at time 0. Notice that the expected value only depends on the realization of the Markov chain controlling the VAR coefficients up to time t, $\xi_1^{\Phi}...\xi_t^{\Phi}$. Let's define the $nm^{\Phi} \times 1$ column vector $q_t \equiv \left[q_t^{1\prime}, ..., q_t^{m^{\Phi}\prime}\right]'$ where $q_t^i = \mathbb{E}_0\left(Z_t 1_{\xi_t^{\Phi}=i}\right) = \mathbb{E}\left(Z_t 1_{\xi_t^{\Phi}=i}|\mathbb{I}_0\right)$ and $1_{\xi_t^{\Phi}=i}$ is an indicator variable that is one when regime i is in place. Note that:

$$q_t^i = \mathbb{E}_0\left(Z_t 1_{\xi_t=i}\right) = \mathbb{E}_0\left(Z_t|\xi_t = i\right)\pi_t^i$$

where $\pi_t^{\Phi,i} = P_0\left(\xi_t^{\Phi} = i\right) = P\left(\xi_t^{\Phi} = i|\mathbb{I}_0\right)$. Therefore we can obtain the conditional expectation $\mathbb{E}_0(Z_t)$ as $\mathbb{E}_0(Z_t) = \sum_{i=1}^{m} q_t^i = wq_t$, where the matrix $w = [I_n, ..., I_n]$ is obtained placing side by side m^{Φ} n-dimensional identity matrices. This is a convenient result because while the law of motion of Z_t is not Markov, the law of motion of q_t is. Following Costa et al. (2004), Bianchi (2016) shows that the law of motion of q_t is given by:

$$q_t = C\pi_t^{\Phi} + \Omega q_{t-1};\ \pi_t^{\Phi} = H^{\Phi}\pi_{t-1}^{\Phi} \tag{11}$$

with $\Omega = bdiag\left(A_1, ..., A_{m^{\Phi}}\right)\left(H^{\Phi} \otimes I_n\right)$ and $C = bdiag\left(c_1, ..., c_{m^{\Phi}}\right)$, where $\otimes$ represents the Kronecker product and $bdiag$ is a matrix operator that takes a sequence of matrices and uses them to construct a block diagonal matrix.

Under the assumption of mean square stability, the process for q_t converges to finite values. Then, given a sequence of probabilities $\pi^{\Phi,T}$ or a posterior draw for the regime sequence ξ^T, the discount-rate news and cash-flow news can be computed as (see Appendix E for a proof):

$$N_{DR}^T = e_1'w\left[\lambda^q v^{q,T} + \lambda^{\pi} v^{\pi,T}\right] \tag{12}$$
$$N_{CF}^T = e_1'w\left[\left(I_r + \lambda^q\right)v^{q,T} + \lambda^{\pi} v^{\pi,T}\right] \tag{13}$$
$$u^T = e_1'wv^{q,T} \tag{14}$$

where $\lambda^q = \left(I_{nm^{\Phi}} - \rho\Omega\right)^{-1}\rho\Omega$, $\lambda^{\pi} = \left(I_{nm^{\Phi}} - \rho\Omega\right)^{-1}\rho CH\left(I_r - \rho H\right)^{-1}$, $v_t^q = q_{t+1|t+1} - q_{t+1|t}$, and $v_{t+1}^{\pi} = \pi_{t+1|t+1}^{\Phi} - \pi_{t+1|t}^{\Phi}$, where $\pi_{t|t}^{\Phi}$ is a column vector whose i-th element coincides with $\pi_{t|t}^{\Phi,i} = P_t\left(\xi_t^{\Phi} = i\right)$, the probability of being in regime i at time t conditional on the information set available at time t. It is worth emphasizing that news now has two components. The first one is represented by the standard Gaussian innovation, while the second component derives from the revision in beliefs about the regime that is in place: $v_{t+1}^{\pi} = \pi_{t+1|t+1}^{\Phi} - \pi_{t+1|t}^{\Phi}$. When the two regimes coincide, formulas (12)-(14) collapse to (9). Therefore, the above formulas can be treated as a generalization of the ones used in Campbell and Vuolteenaho (2004).

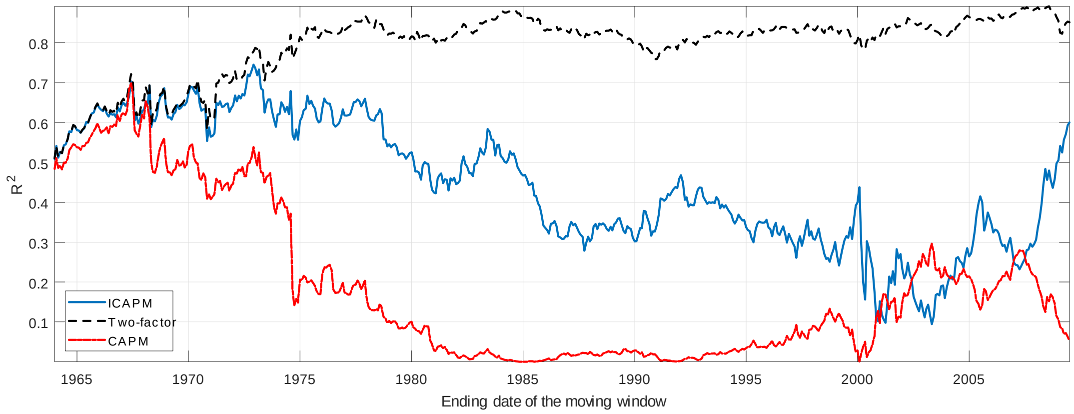

Figure 11: **Explanatory power of the ICAPM over moving windows.** The figure reports the explanatory power as measured by the R^2 of three models: The unrestricted two-factor model, the Intertemporal CAPM, and the traditional CAPM. The betas and average returns are computed over moving windows of 35 years. The horizontal axis reports the ending date of the rolling window. For example 1965 corresponds to the sample February 1930-January 1965. The dependent variables are the average returns of the 25 Fama-French portfolios.

For the practical implementation of the formulas presented above, the vector of regime probabilities and parameters need to be replaced by their corresponding estimates. In the benchmark results presented below, I use the parameter estimates obtained using the entire sample and the corresponding filtered probabilities. An alternative approach would be to assume that agents in the economy acts as econometricians and estimate the model recursively. In this second case, the agents' information set and the econometrician's information set are aligned (up to revision in the data). Results for this second approach are very similar and are described in Appendix F.

It is worth pointing out that the approach described above can model situations in which not all m regimes are stable. This is because in order to be able to compute the news, we only need the discounted expectations to be stable. Mean square stability guarantees stability for first and second moments, i.e., covariance stationarity. Notice that this is in fact more than what is necessary for two reasons. First, the VAR implementation does not require the variance to be stable, but only that agents' expectations converge. Second, even if first moments are not stable, *discounted* first moments might be. However, it might be argued that imposing covariance stationarity is still desirable, given that it implies that agents' uncertainty converges to a finite value no matter the regime that is in place today. For the estimates considered in this paper, both regimes were determined to be stable.

6.3 Evolution of the explanatory power of the models

Figure 11 reports the evolution of R^2 for the three models over rolling windows of 35 years.[5] The dependent variables are the average returns of the 25 Fama-French portfolios over the same time period. I drop the extreme small-growth portfolio that is often found to be an outlier in asset-pricing models. The explanatory power of all models is very high at the beginning of the sample, and initially it tends to increase as the window moves to the right. However, past the 1970s the performance of the CAPM starts to quickly deteriorate, with a very visible drop around 1975. On the contrary, the performance of the unrestricted two-factor model remains substantially high, with values often above 80%. However, this model does not impose economically motivated restrictions on the premia, so it is not surprising that it delivers a higher R^2. The ICAPM does very well until the mid-1980s, even if its performance starts following a downward trend. By the mid-1990s, the R^2 starts fluctuating around 30%, very far from the 60% attained during the first half of the sample. However, as the window approaches the most recent financial crisis, the explanatory power of the ICAPM increases steeply and the R^2 touches 60%. This is a remarkable improvement in fit given that the last time that the ICAPM explanatory power crossed the 60% threshold was toward the end of 1978, and it has not been larger than 50% since the first half of 1985. Instead, the performance of the CAPM does not show any significant recovery. Appendix H shows that similar results hold using expanding windows as opposed to recursive windows: As long as the Great Depression or the Great Recession (or both) are included in the sample, the ICAPM model is able to explain the cross section of asset returns.

These results have some suggestive implications. First of all, they highlight the role played by the Great Depression and the Great Recession. Once these events are included in the analysis, the ICAPM performance substantially improves. Furthermore, the fact that the performance of the ICAPM improves, while the explanatory power of the CAPM remains unsatisfactory, implies that distinguishing between the two sources of risk is crucial and that this distinction becomes particularly meaningful in the aftermath of exceptional events.

To understand what drives the improvement in fit of the ICAPM, Figure 12 reports betas and deviations of portfolio returns from their predicted values for the 25 Fama-French portfolios. The first and second columns report cash-flow betas and discount-rate betas. The third column reports the composition of the market beta. This is computed as the ratio between the cash-flow beta and the sum of cash-flow and discount-rate betas. Recall that the market beta is obtained summing the two betas. Finally, the fourth column contains the deviations of portfolio returns with respect to the values predicted by the ICAPM. In each row, the solid blue lines refer to the portfolios indicated on the vertical axis of the figure.

A series of interesting patterns emerge. First, as in Campbell and Vuolteenaho (2004), over

[5] The R^2 is computed as 1 □ RSS/RSM where RSS is the residual sum of squares and RSM is the residual sum of squares when only the constant is used as a regressor. Note that in the ICAPM this variable can become negative because the model imposes a restriction on the premia of the discount-rate beta.

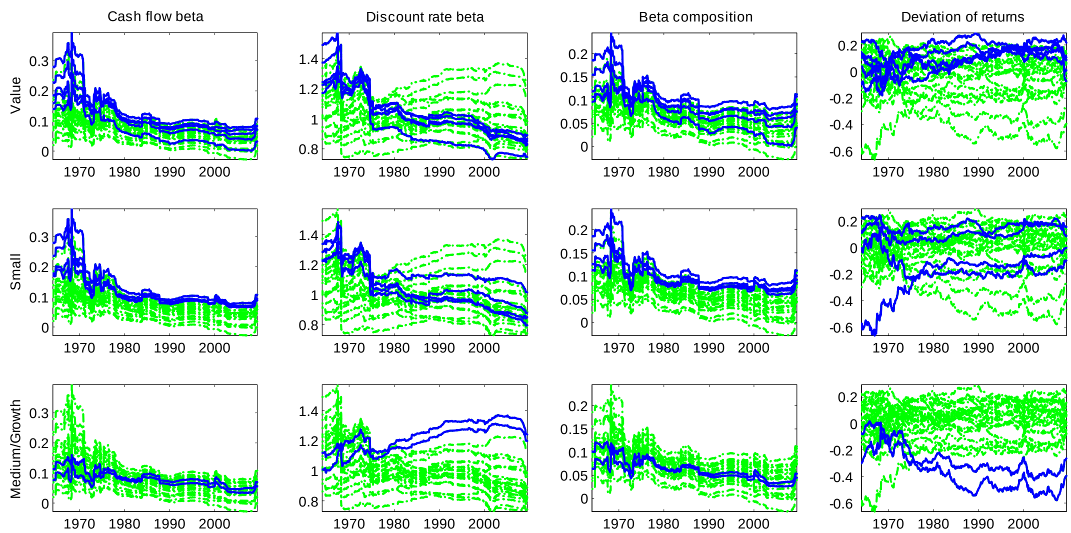

Figure 12: **Betas and predicted returns.** The first and second columns report cash flow betas and discount rate betas for the 25 Fama-French portfolios. The third column reports the composition of the market beta, and it is computed as the ratio between the cash-flow beta and the sum of cash-flow and discount-rate betas. Finally, the fourth column contains the deviations of portfolio returns from the values predicted by the Intertemporal CAPM. In each row, the solid blue lines refer to the portfolios indicated on the vertical axis of the first column.

time, the value and small stocks experience a pronounced decline in the market beta with respect to the other portfolios. However, this decline is mostly driven by a fall in their discount rate betas, while their cash-flow betas remain on the upper side of the spectrum. This pattern implies a change in the composition of the market beta that in turn explains the success of the ICAPM over the CAPM. The ICAPM separates the different sources of risk associated with the two betas. Second, the most notable deviations of stock market returns from what is predicted by the ICAPM are caused by two medium/growth portfolios.[6] When analyzing the period antecedent the current crisis and excluding the Great Depression, these portfolios have stock market returns that are too low with respect to what is predicted by the ICAPM. While these stocks show a relatively large increase in their discount rate beta and a stable composition for the beta, their average returns do not adequately reflect such an increase in risk. Finally, this anomaly is largely reduced toward the end of the sample, and at the same time, the returns of the small and value portfolios also move closer to their predicted returns.

From these results, we can infer that in order to adequately price the cross section of asset returns it is important to be able to observe the behavior of the assets during exceptional events such as the Great Recession. The relative performance of the different portfolios change substantially during these events. It is also important to emphasize that this is not the result of drastic changes in the betas. Even if we observe a partial increase in the cash-flow betas during the late years, the relative ranking of the portfolios with respect to the betas appears quite stable. It is

[6]These two portfolios correspond to the ones with the second and third smallest market value among the five growth portfolios.

therefore the change in the relative performance of the different portfolios during a time of distress that is largely responsible for the improvement in fit. In order to formalize this point, I regressed the average returns over the last window of time (June 1974-May 2009) on the betas computed using the window of time right before Bear Stearns received a loan from the Federal Reserve Bank of New York (April 1975-March 2008). The resulting R^2 is still high, 53.28%, even if lower than the value obtained aligning betas and average returns.

There are several possible explanations for why value stocks might perform worse during financial crises. Zhang (2005) argues that the value premium arises naturally in a neoclassical model because of costly reversibility and countercyclical price of risk. During bad times firms would find it optimal to disinvest, implying that assets in place are riskier than growth options. It seems reasonable that this distinction becomes particularly relevant during financial crises. Furthermore, almost by definition, value stocks include firms that markets believe might have less prospects of growth in the future. While this is not necessarily a problem during regular times, it can become a serious issue when credit availability is limited, real activity is low, and the price of risk is high. Similarly, Campbell and Vuolteenaho (2004) suggest that during the Great Depression and in its aftermath, value stocks might include a significant fraction of *fallen angels* that accumulated large amounts of debt during the crisis and were therefore inherently riskier.

The Great Depression regime also plays another key role: it shapes agents' expectations. In order to address the importance of this channel, I reconsider the evolution of the explanatory power of the ICAPM under three different assumptions about the way agents form expectations. Under the benchmark model, agents take into account the possibility of regime changes. This corresponds to the benchmark case. In the second case, agents form expectations according to the anticipated utility assumption. This implies that the probability assigned to the two regimes are not moving over time. Under this assumption, the series for the news are computed by replacing the estimated transition matrix H^{Φ} with the identity matrix in the formulas presented in Subsection 6.2. In the last scenario, agents form expectations based on a fixed coefficients VAR estimated over the whole sample. Thus, agents use the information of the Great Depression and the Great Recession, but without recognizing that these are exceptional events.

Figure 13 presents the results. The blue solid line corresponds to the benchmark case, the black dashed line reports the results for the case of anticipated utility, the red dashed-dotted line corresponds to the fixed coefficients VAR case. It is interesting to note that three cases return a similar fit over the early subsamples. The benchmark model and the anticipated case start diverging in the mid-1970s. It is worth recalling that over the very same months the CAPM also had a drastic decline in the fit (see Figure 11). In other words, exactly when the distinction between the CAPM and the ICAPM becomes more meaningful, we observe a discrete drop in the fit of the ICAPM under the assumption of anticipated utility. The anticipated utility assumption becomes relatively more innocuous toward the end of the sample. This seems sensible given that this is the period of time during which the dynamics resembling the Great Depression present

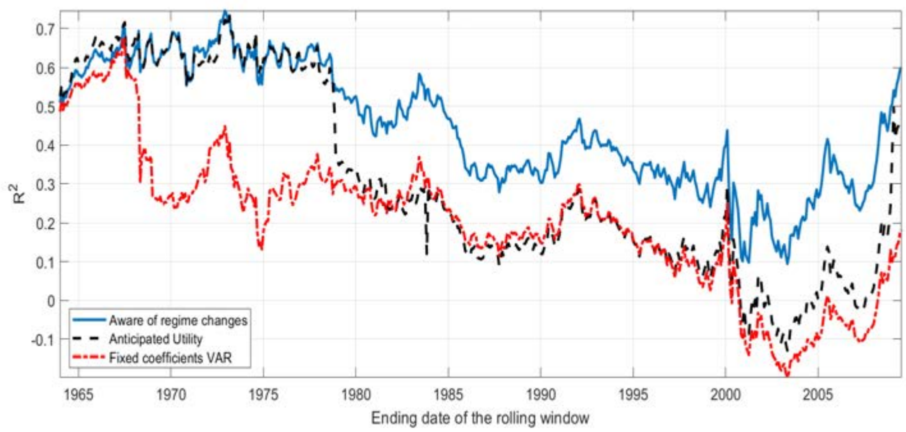

Figure 13: **The role of agents' beliefs.** The figure reports the explanatory power as measured by the R^2 for the ICAPM under three different assumptions about the way agents form expectations. The blue solid line corresponds to the benchmark case in which agents take into account the possibility of regime changes. The black dashed line corresponds to the case in which agents form expectations according to the anticipated utility assumption. In this second case, the probability assigned to the two regimes are not moving over time. The third case (red dashed-dotted line) assumes a fixed coefficients VAR. The betas and average returns are computed over moving windows of 35 years. The horizontal axis reports the ending date of the rolling window. For example, 1965 corresponds to the sample February 1930-January 1965. The dependent variables are the average returns of the 25 Fama-French portfolios.

themselves. However, the gap in the explanatory power is still approximately 15%. Finally, the R^2 of the ICAPM under the benchmark case with regime changes is always higher than the case with fixed coefficients. The difference in R^2 is very large, always positive, 24.30% in average, and can be as large as 44.84%. Thus, even if in the fixed coefficients VAR case agents use the information of the Great Depression when forming expectations, the fact that such information is mixed together with the dynamics during regular times leads to a large decline in the explanatory power of the ICAPM.

In summary, the Great Depression regime also plays a key role in accounting for the cross section of asset returns during regular times because it shapes the way agents form expectations. In fact, during regular times, it becomes particularly important to take into account the possibility of regime changes because no exceptional events are present over the sample. This is not enough to completely compensate for the fact that no financial crisis is observed, but it still determines an improvement in the fit of the ICAPM.

7 Conclusions

Using an MS-VAR, I have identified a Great Depression regime and shown that its probability has been close to zero until the most recent recession. In February 2009, the probability of the Great Depression regime spiked to cross 50%, and it was larger than 80% when using real time estimates. During the early months of both the Great Depression and the Great Recession, the Value spread was increasing while the stock market was falling. However, during the Great Recession, this pattern eventually reverted, and the probability of the Great Depression regime experienced a sharp drop, arguably in response to robust government interventions, signaling that the U.S. was in fact able to avoid a financial meltdown. To substantiate this argument, I show that a parsimonious extension of Gertler and Karadi (2011) can account for the behavior of real activity and asset valuation observed during the two events.

I then argue that the Great Recession and the Great Depression were not like any other financial crises. They were both global phenomena that originated in the United States and had severe consequences for all developed economies that were affected. While the US economy recovered fairly quickly during the Great Recession, the same it is not true for many European economies. Finally, the pattern of a stock market decline paired with an increase in the value spread during the Great Recession seems common to other modern economies, suggesting that severe financial do not affect all stocks symmetrically. I formalize this idea by showing that the existence of the Great Depression regime is important to understand the cross section of asset returns.

References

Ang, A. and G. Bekaert (2002). Regime switches in interest rates. *Journal of Business & Economic Statistics 20*(2), 163–82.

Ang, A. and A. Timmermann (2012). Regime Changes and Financial Markets. *Annual Review of Financial Economics 4*(1), 313–337.

Bai, H., K. Hou, H. Kung, and L. Zhang (2015). The CAPM Strikes Back? An Investment Model with Disasters. LBS working paper.

Barro, R. (2009). Rare disasters, asset prices, and welfare costs. *The American Economic Review 99*(1), 243–264.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics 121*, 823–866.

Bernanke, B. S. (1983). Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression. *American Economic Review 73*, 257–276.

Bianchi, F. (2016). Methods for Measuring Expectations and Uncertainty in Markov-switching Models. *Journal of Econometrics 190*(1), 79 – 99.

Bianchi, F., M. Lettau, and S. C. Ludvigson (2016). Monetary Policy and Asset Valuation: Evidence From a Markov-Switching cay. NBER Working Paper No. 22572.

Bollerslev, T. and V. Todorov (2011). Tails, Fears, and Risk Premia. *Journal of Finance 66*(6), 2165–2211.

Campbell, J. Y. (1991). A Variance Decomposition for Stock Returns. *Economic Journal 101*(405), 157–179.

Campbell, J. Y. (1993). Intertemporal Asset Pricing Without Consumption Data. *American Economic Review 83*, 487–512.

Campbell, J. Y., S. Giglio, and C. Polk (2013). Hard times. *Review of Asset Pricing Studies 3*(1), 95–132.

Campbell, J. Y., S. Giglio, C. Polk, and R. Turley (2014). An intertemporal capm with stochastic volatility. Harvard working Paper.

Campbell, J. Y. and R. J. Shiller (1988). The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. *Review of Financial Studies 1*, 195–228.

Campbell, J. Y. and T. Vuolteenaho (2004). Bad Beta, Good Beta. *American Economic Review 94*, 1249–1275.

Costa, O., M. Fragoso, and R. Marques (2004). *Discrete-Time Markov Jump Linear Systems.* New York: Springer.

Eichengreen, B. and K. O'rourke (2010). A Tale of Two Depressions: What do the new data tell us? VoxEU article.

Epstein, L. G. and S. E. Zin (1989). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica 57*(4), 937–69.

Fama, E. F. and K. R. French (1992). The Cross-Section of Expected Stock Returns. *Journal of Finance 47*(2), 427–65.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Farmer, R. E., D. F. Waggoner, and T. Zha (2009). Understanding Markov-Switching Rational Expectations Models. *Journal of Economic Theory 144*, 1849–1867.

Foerster, A. T. (2015). Financial crises, unconventional monetary policy exit strategies, and agentsŒş expectations. *Journal of Monetary Economics 76*(C), 191–207.

Gabaix, X. (2012). Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance. *Quarterly Journal of Economics 127*(2), 645–700.

Gertler, M. and P. Karadi (2011). A Model of Unconventional Monetary Policy. *Journal of Monetary Economics 58*(1), 17–34.

Geweke, J. (1992). Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments. *Bayesian Statistics 4*, 169–193.

Gourio, F. (2012, October). Disaster Risk and Business Cycles. *American Economic Review 102*(6), 2734–66.

Gulen, H., Y. Xing, and L. Zhang (2011). Value versus Growth: Time-Varying Expected Stock Returns. *Financial Management 40*(2), 381–407.

Jermann, U. and V. Quadrini (2012). Macroeconomic Effects of Financial Shocks. *American Economic Review 102*(1), 238–271.

Jorda, O., M. Schularick, and A. Taylor (2016, September). Macrofinancial History and the New Business Cycle Facts. In *NBER Macroeconomics Annual 2016, Volume 31*, NBER Chapters, pp. 213–263. National Bureau of Economic Research, Inc.

Julliard, C. and A. Ghosh (2012). Can Rare Events Explain the Equity Premium Puzzle? *Review of Financial Studies 25*(10), 3037–3076.

Kim, C.-J. and C. R. Nelson (1999). *State-Space Models with Regime Switching.* Cambridge, Massachusetts: MIT Press.

Krugman, P. R. (2009). Fighting Off Depression. New York Times OP-ED, January 4, 2009.

Lettau, M., S. C. Ludvigson, and J. A. Wachter (2008). The Declining Equity Premium: What Role Does Macroeconomic Risk Play? *Review of Financial Studies 21*(4), 1653–1687.

Merton, R. (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica 41*(4), 867–887.

Nakamura, E., J. Steinsson, R. Barro, and J. Ursua (2013). Crises and Recoveries in an Empirical Model of Consumption Disasters. *American Economic Journal: Macroeconomics 5*(3), 35–74.

Pesaran, M. H., D. Pettenuzzo, and A. Timmermann (2006). Forecasting Time Series Subject to Multiple Structural Breaks. *Review of Economic Studies 73*, 1057–1084.

Raftery, A. and S. Lewis (1992). How many iterations in the Gibbs sampler? *Bayesian Statistics 4*, 763–773.

Rietz, T. A. (1988). The Equity Risk Premium: A Solution. *Journal of Monetary Economics 22*, 117–131.

Shiller, R. J. (2000). *Irrational exuberance.* Princeton University Press.

Sims, C. A. and T. Zha (1998). Bayesian Methods for Dynamic Multivariate Models. *International Economic Review 39*(4), 949–968.

Sims, C. A. and T. Zha (2006). Were There Regime Switches in US Monetary Policy? *American Economic Review 91*(1), 54–81.

Solomon, D. (2009). Market Pans Bank Rescue Plan. Wall Street Journal, 11 February 2009.

Wachter, J. A. (2013, 06). Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility? *Journal of Finance 68*(3), 987–1035.

Zhang, L. (2005, 02). The Value Premium. *Journal of Finance 60*(1), 67–103.

A Priors

Table 3 describes the priors used for the estimation of the MS-VAR. The priors are very loose and symmetric across regimes. Below, I describe more in detail how they have been obtained. The assumption of covariance stationarity implies a truncation of the priors as described in the table. The truncated prior is implemented by dropping the draws that imply non-stationarity. However, in the estimates the constraint implied by the truncation of the priors is rarely binding.

The priors for the VAR coefficients and the covariance matrix are symmetric across regimes and are obtained running univariate autoregressions for each endogenous variable:

$$y_{i,t} = a_i y_{i,t-1} + v_t \sigma_i$$

The prior for the VAR coefficients is:

$$B = vec\left(\Phi_{\xi_t^\Phi}\right) \sim norm\left(B_0, S_0 \otimes N_0^{-1}\right)$$

The autoregressive elements of B_0 are equal to the AR(1) coefficients, while all the other elements are set to zero. As in Sims and Zha (1998), the variance of the prior distribution is specified by a number of hyperparameters that pin down N_0. The choice of hyperparameters implies a fairly loose prior for the VAR coefficients. Let λ be a (5×1) vector containing the hyperparameters. The diagonal elements of N_0^{-1} corresponding to autoregressive coefficients are given as $\left(\frac{\lambda_0 \lambda_1}{\sigma_i l^{\lambda_3}}\right)^2$ where σ_j denotes the variance of the error from the AR regression for the jth variable and $l = 1...L$ denotes the lags in the VAR ($L = 1$ in the models considered in this paper). The intercept terms in N_0^{-1} are controlled by the term $(\lambda_0 \lambda_4)^2$. The choice for the hyperparameters are $\lambda_0 = 1$, $\lambda_1 = 1$, $\lambda_2 = 1$, $\lambda_3 = 0.5$ and $\lambda_4 = 1$ and $S_0 = V_0 diag\left(\{\sigma_i^2\}_{i=1...n}\right)$, with $V_0 = 9$. The priors for the covariance matrices are symmetric across regimes and described by an inverse Wishart distribution with mean $S_0 = V_0 diag\left(\{\sigma_i^2\}_{i=1...n}\right)$, with $V_0 = 9$: $\Sigma_{\xi_t^\Sigma} \sim IW(S_0, V_0)$.

Each column of H^Φ and H^Σ is modeled according to a Dirichlet distribution whose properties are described in Table 3: $H^s(\cdot, i) \sim D(a_{ii}^s, a_{ij}^s)$, $s = \Phi, \Sigma$. I choose $a_{ii}^\Sigma = 10$, $a_{ij}^\Sigma = 2$, $a_{ii}^\Phi = 80$, $a_{ij}^\Phi = 2$. Note that the priors for the transition matrices are symmetric across regimes. I also estimated looser priors for the transition matrices and obtained very similar results.

B Likelihood and regime probabilities

Define the combined regime $\xi_t \equiv \left(\xi_t^\Phi, \xi_t^\Sigma\right)$, the associated transition matrix $H \equiv H^\Phi \otimes H^\Sigma$, and vector $\theta_{\xi_t} \equiv \left(\Phi_{\xi_t^\Phi}, \Sigma_{\xi_t^\Sigma}\right)$ with the corresponding set of parameters. For each draw of the parameters θ_{ξ_t} and H, we can then compute the filtered probabilities $\pi_{t|t}$, or smoothed probabilities $\pi_{t|T}$, of the regimes conditional on the model parameters. The filtered probabilities reflect the probability

$\xi_t^{\Phi}=1,2$	ER_t	TY_t	PE_t	VS_t	$const$
ER_{t+1}	0.1236 (□0.0857,0.3316)	0 (□0.0216,0.0216)	0 (□0.1357,0.1357)	0 (□0.1552,0.1552)	0.0049 (□0.0331,0.0429)
TY_{t+1}	0 (□1.3427,1.3427)	0.9472 (0.8084,1.0866)	0 (□0.8765,0.8765)	0 (□1.0012,1.0012)	0.0823 (□0.1620,0.3265)
PE_{t+1}	0 (□0.2140,0.2140)	0 (□0.0221,0.0221)	0.9888 (0.8497,1.1282)	0 (□0.1593,0.1593)	0.0324 (□0.0066,0.0712)
VS_{t+1}	0 (□0.1868,0.1868)	0 (□0.0195,0.0195)	0 (□0.1218,0.1218)	0.9909 (0.8514,1.1303)	0.0147 (□0.0193,0.0488)

$\xi_t^{\Sigma}=1,2,3$	u_{ER}	u_{TY}	u_{PE}	u_{VS}
u_{ER}	0.0546 (0.0797,0.1981)	0 (□0.0610,0.0610)	0 (□0.0097,0.0097)	0 (□0.0085,0.0085)
u_{TY}	0 (□0.5660,0.5660)	0.3515 (0.5124,1.2733)	0 (□0.0624,0.0624)	0 (□0.0545,0.0545)
u_{PE}	0 (□0.5660,0.5660)	0 (□0.5660,0.5600)	0.0559 (0.0814,0.2024)	0 (□0.0086,0.0086)
u_{VS}	0 (□0.5660,0.5660)	0 (□0.5660,0.5660)	0 (□0.5660,0.5660)	0.0489 (0.0714,0.1773)

H^{Φ}	$\xi_t^{\Phi}=1$	$\xi_t^{\Phi}=2$
$\xi_{t+1}^{\Phi}=1$	0.9875 (0.9599,0.9912)	0.0125 (0.0088,0.0401)
$\xi_{t+1}^{\Phi}=2$	0.0125 (0.0088,0.0401)	0.9875 (0.9599,0.9912)

H^{Σ}	$\xi_t^{\Sigma}=1$	$\xi_t^{\Sigma}=2$	$\xi_t^{\Sigma}=3$
$\xi_{t+1}^{\Sigma}=1$	0.8182 (0.5954,0.8327)	0.0909 (0.0553,0.2315)	0.0909 (0.0553,0.2315)
$\xi_{t+1}^{\Sigma}=2$	0.0909 (0.0553,0.2315)	0.8182 (0.5952,0.8324)	0.0909 (0.0553,0.2315)
$\xi_{t+1}^{\Sigma}=3$	0.0909 (0.0553,0.2315)	0.0909 (0.0553,0.2315)	0.8182 (0.5955,0.8324)

Table 3: Priors for the parameters. The three sets of tables contain modes and 68% error bands for the priors of the parameters of the Markov-switching VAR. The priors are obtained running univariate autoregressions for each of the variables in the model, and they are symmetric across regimes.

of a regime conditional on the data up to time t, $\pi_{t|t} = p(\xi_t|Y^t; H, \theta_{\xi_t})$, for $t = 1, ..., T$, and are part of the output obtained computing the likelihood function associated with the parameter draw H, θ_{ξ_t}. The filtered probabilities can be obtained using the following recursive algorithm:

$$\pi_{t|t} = \frac{\pi_{t|t□1} \odot \eta_t}{\mathbf{1}' \, □\pi_{t|t□1} \odot \eta_t)} \quad (15)$$

$$\pi_{t+1|t} = H\pi_{t|t} \quad (16)$$

$$p(Z_t|Z^{t□1}) = \mathbf{1}' \, □\pi_{t|t□1} \odot \eta_t) \quad (17)$$

where η_t is a vector whose jth element contains the conditional density $p(Z_t|\xi_t = i, Z^{t□1}; H, \theta_{\xi_t})$, the symbol $\odot$ denotes element by element multiplication, and $\mathbf{1}$ is a vector with all elements equal to 1. To initialize the recursive calculation, we need an assumption on the distribution of ξ_0. We assume that the six regimes have equal probabilities $p(\xi_0 = i) = 1/6$ for $i = 1...m$. The likelihood for the entire data sequence Z^T is obtained multiplying the one-step-ahead conditional likelihoods $p(Z_t|Z^{t□1})$:

$$p\,□Z^T|\theta) = \prod_{t=1}^{T} p\,□Z_t|Z^{t□1})$$

The smoothed probabilities reflect all the information that can be extracted from the whole data sample, $\pi_{t|T} = p(\xi_t|Z^T; H, \theta_{\xi_t})$. The final term $\pi_{T|T}$ is returned with the final step of the filtering algorithm. Then a recursive algorithm can be implemented to derive the other probabilities:

$$\pi_{t|T} = \pi_{t|t} \odot \left[H' \left(\pi_{t+1|T} \, (\div) \, \pi_{t+1|t}\right)\right]$$

where $(\div)$ denotes element by element division.

Finally, it is possible to obtain the filtered and smoothed probabilities for each of the two independent chains by integrating out the other chain. For example, if we are interested in $\pi^{\Phi}_{t|t} = p(\xi^{\Phi}_t | Y^t; H, \theta_{\xi_t})$ we have:

$$\pi^{\Phi,i}_{t|t} = p(\xi^{\Phi}_t = i | Y^t; H, \theta_{\xi_t}) = \sum_{j=1}^{m} p(\xi_t = \{i,j\} | Y^t; H, \theta_{\xi_t})$$

Similarly, the smoothed probabilities are obtained as:

$$\pi^{\Phi,i}_{t|T} = p(\xi^{\Phi}_t = i | Y^T; H, \theta_{\xi_t}) = \sum_{j=1}^{m} p(\xi_t = \{i,j\} | Y^T; H, \theta_{\xi_t}).$$

C Gibbs sampling algorithm

Both the VAR coefficients and the covariance matrix can switch and the regimes are assumed to be independent. Draws for the parameters of the model can be made following the following Gibbs sampling algorithm:

1. Sampling ξ^{Φ}_t and ξ^{Σ}_t given $\Phi_{\xi^{\Phi}_t}, \Sigma_{\xi^{\Sigma}_t}, H^{\Phi}, H^{\Sigma}$: Following Kim and Nelson (1999) I use a Multi-Move Gibbs sampling to draw ξ^{Φ}_t from $f(\xi^{\Phi}_t | Z^T, \Phi_{\xi^{\Phi}_t}, \Sigma_{\xi^{\Sigma}_t}, H^{\Phi}, H^{\Sigma}, \xi^{\Sigma}_t)$ and ξ^{Σ}_t from $f(\xi^{\Sigma}_t | Z^T, \Phi_{\xi^{\Phi}_t}, \Sigma_{\xi^{\Sigma}_t}, H^{\Phi}, H^{\Sigma}, \xi^{\Phi}_t)$.

2. Sampling $\Sigma_{\xi^{\Sigma}_t}$ given $\Phi_{\xi^{\Phi}_t}, \xi^{\Phi}_t, \xi^{\Sigma}_t$: Given $\Phi_{\xi^{\Phi}_t}$ and $\xi^{\Phi,T}$, we can compute the residuals of the MS-VAR at each point in time. Then, given ξ^{Σ}_t, we can group all the residuals that pertain to a particular regime. Therefore, $\Sigma_{\xi^{\Sigma}_t}$ can be drawn from an inverse Wishart distribution for $\xi^{\Sigma}_t = 1...m^{\Sigma}$.

3. Sampling $\Phi_{\xi^{\Phi}_t}$ given $\Sigma_{\xi^{\Sigma}_t}, \xi^{\Phi}_t, \xi^{\Sigma}_t$: When drawing the VAR coefficients, we need to take into account the heteroskedasticity implied by the switches in $\Sigma_{\xi^{\Sigma}_t}$. This can be done following the following steps for each $i = 1...m^{\Phi}$:

   (a) Based on $\xi^{\Phi,T}$, collect all the observation such that $\xi^{\Phi}_t = i$.

   (b) Divide the data that refer to $\xi^{\Sigma}_t = j$ based on $\xi^{\Sigma,T}$. We now have a series of subsamples for which VAR coefficients and covariance matrices are fixed: $\left(\xi^{\Phi}_t = i,\ \xi^{\Sigma}_t = 1\right)$, ..., $\left(\xi^{\Phi}_t = i,\ \xi^{\Sigma}_t = m^{\Sigma}\right)$. Denote these subsamples with $\left(y_{i,\xi^{\Sigma}_t},\ x_{i,\xi^{\Sigma}_t}\right)$, where the y_{i,ξ^{Σ}_t} and x_{i,ξ^{Σ}_t} denote left-hand-side and right-hand-side variables in the MS-VAR. Notice that some of these subsamples might be empty.

(c) Apply recursively the formulas for the posterior of VAR coefficients conditional on a known covariance matrix. Therefore, for $j = 1...m^{\Sigma}$ the following formulas need to be applied recursively:

$$
\begin{aligned}
P_T^{-1} &= P_L^{-1} + \Sigma_{\xi_t^{\Sigma}}^{-1} \otimes (x'_{i,\xi_t^{\Sigma}} x_{i,\xi_t^{\Sigma}}) \\
B_T &= B_L + (\Sigma_{\xi_t^{\Sigma}}^{-1} \otimes x'_{i,\xi_t^{\Sigma}}) vec(y_{i,\xi_t^{\Sigma}}) \\
P_L^{-1} &= P_T^{-1}, B_L = B_T
\end{aligned}
$$

where the algorithm is initialized using the priors for the VAR coefficients $B_L = B_0$ and $P_L^{-1} = P_0^{-1} = \left(S_0 \otimes N_0^{-1}\right)^{-1}$. Notice that this implies that if there are not any observations for a particular regime, then the posterior will coincide with the priors. With proper priors, this is not a problem.

(d) Make a draw for the VAR coefficients $vec\left(\Phi_{\xi_t^{\Phi}}\right) \sim N\left(P_T B_T, P_T\right)$ with $\xi_t^{\Phi} = i$.

4. Sampling H^{Φ} and H^{Σ}: Given the draws for the state variables $\xi^{\Phi,T}$ and $\xi^{\Sigma,T}$, the transition probabilities are independent of Y_t and the other parameters of the model and have a Dirichlet distribution. For each column of H^{Φ} and H^{Σ}, the posterior distribution is given by:

$$
H^s(:,i) \sim D(a_{ii}^s + \eta_{ii}^s, a_{ij}^s + \eta_{ij}^s), \;\; s = \Phi, \Sigma
$$

where η_{ij}^{Φ} and η_{ij}^{Σ} denote respectively the numbers of transitions from state i^{Φ} to state j^{Φ} and from state i^{Σ} to state j^{Σ}.

D Properties of the regimes

Figure 14 reports the distribution for the difference between the parameter of the VAR coefficients. Figure 15 reports the distribution for the difference between the elements of the covariance matrix under the Term Yield volatility regime and the Low volatility regime. Finally, Figure 16 reports the difference between the elements of the covariance matrix under the Term Yield volatility regime and the High volatility regime. These are computed taking the difference between the corresponding parameters for each draw from the Gibbs sampling algorithm.

E Cash-flow and discount-rate news with regime changes

Consider an MS-VAR:

$$
Z_t = c_{\xi_t} + A_{\xi_t} Z_{t-1} + R_{\xi_t} \Sigma_{\xi_t} \varepsilon_t
$$

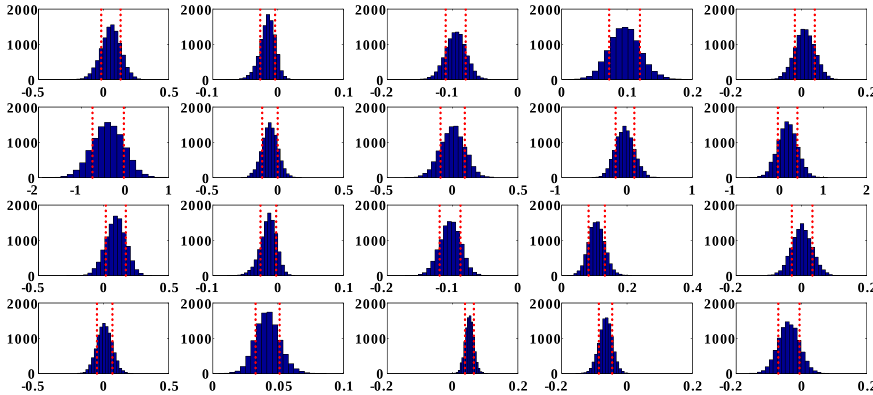

Figure 14: The figure contains histograms and 68% error bands for the pairwise differences of the VAR coefficients across the two regimes. This can be regarded as a "test" for the null hypothesis that the two parameters are the same across the two regimes.

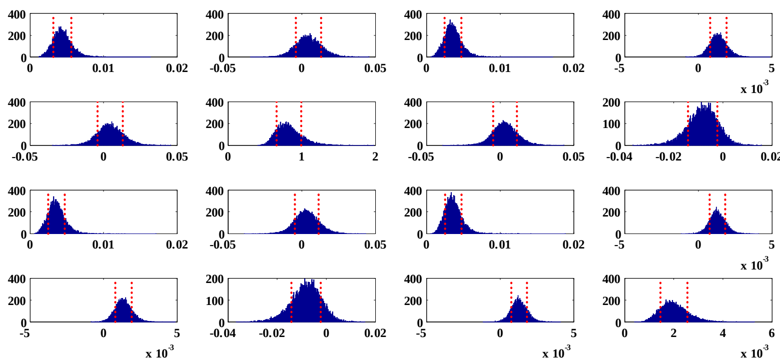

Figure 15: The figure contains histograms and 68% error bands for the pairwise differences of the covariance matrix under the Term Yield Volatility regime and the Low volatility regime. This can be regarded as a "test" for the null hypothesis that the two parameters are the same across the two regimes.

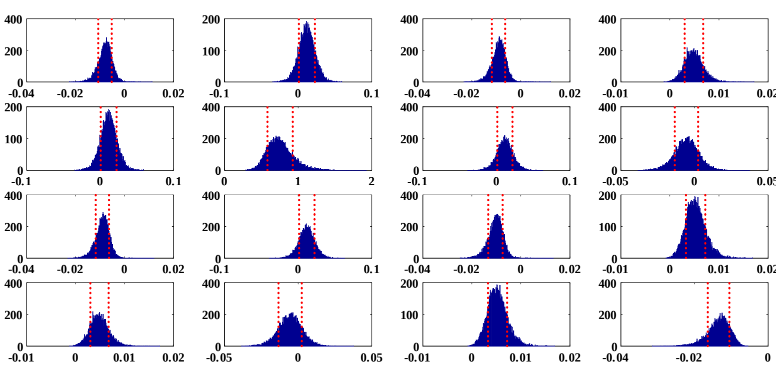

Figure 16: The figure contains histograms and 68% error bands for the pairwise differences of the covariance matrix under the Term Yield Volatility regime and the High volatility regime. This can be regarded as a "test" for the null hypothesis that the two parameters are the same across the two regimes.

where Z_t is a column vector containing n variables observable at time t and $\xi_t = 1, ..., m$, with m the number of regimes, evolves following the transition matrix H.

Define the column vectors q_t and π_t:

$$q_t = \left[q_t^{1\prime}, ..., q_t^{m\prime}\right]', q_t^i = \mathbb{E}_0\left(Z_t 1_{\xi_t = i}\right), \pi_t = \left[\pi_t^1, ..., \pi_t^m\right]',$$

where $\pi_t^i = P_0(\xi_t = i)$ and $1_{\xi_t = i}$ is an indicator variable that is equal to 1 when regime i is in place and zero otherwise. The law of motion for $\widetilde{q}_t = [q_t', \pi_t']'$ is then given by

$$\underbrace{\begin{bmatrix} q_t \\ \pi_t \end{bmatrix}}_{\widetilde{q}_t} = \underbrace{\begin{bmatrix} \Omega & CH \\ & H \end{bmatrix}}_{\widetilde{\Omega}} \begin{bmatrix} q_{t-1} \\ \pi_{t-1} \end{bmatrix} \tag{18}$$

where $\pi_t = [\pi_{1,t}, ..., \pi_{m,t}]'$, $\Omega = bdiag(A_1, ..., A_m)H$, and $C = bdiag(c_1, ..., c_m)$. Recall that:

$$\mathbb{E}_0(Z_t) = \sum_{i=1}^{m} q_t^i = wq_t, \ w = \left[\underbrace{I_n, ..., I_n}_{m}\right]$$

To compute the news, define:

$$\begin{aligned} q^i_{t+s|t} &= \mathbb{E}_t\left(Z_{t+s} 1_{\xi_{t+s} = i}\right) = \mathbb{E}\left(Z_{t+s} 1_{\xi_{t+s} = i} | \mathbb{I}_t\right) \\ e_1' &= [1, 0, 0, 0]', \ mn = m * n \end{aligned}$$

where $\mathbb{I}_t$ contains all the information that agents have at time t, including the probability of being in one of the m regimes. Note that $q^i_{t|t} = Z_t \pi_t^i$.

Now consider the formula for the discount-rate news:

$$N_{DR,t+1} = (\mathbb{E}_{t+1} - \mathbb{E}_t) \sum_{j=1}^{\infty} \rho^j r_{t+1+j}$$

The first term is:

$$\begin{aligned} \mathbb{E}_{t+1} \sum_{j=1}^{\infty} \rho^j r_{t+1+j} &= \sum_{j=1}^{\infty} \rho^j e_1' w q_{t+1+j|t+1} \\ &= e_1' w \left[\rho q_{t+2|t+1} + \rho^2 q_{t+3|t+1} + \rho^3 q_{t+4|t+1} + ...\right] \\ &= e_1' w (I_r - \rho\Omega)^{-1} \left[\rho\Omega q_{t+1|t+1} + \rho CH (I_r - \rho H)^{-1} \pi_{t+1|t+1}\right] \end{aligned}$$

The second term is:

$$\begin{aligned} \mathbb{E}_t \sum_{j=1}^{\infty} \rho^j r_{t+1+j} &= \sum_{j=1}^{\infty} \rho^j e_1' w q_{t+1+j|t} \\ &= e_1' w (I_r - \rho\Omega)^{-1} \left[\rho\Omega q_{t+1|t} + \rho CH (I_r - \rho H)^{-1} \pi_{t+1|t}\right] \end{aligned}$$

Therefore:

$$\begin{aligned}
N_{DR,t+1} &= e_1' w \left[\lambda^q v_{t+1}^q + \lambda^\pi v_{t+1}^\pi\right] \\
\lambda^q &= (I_r - \rho\Omega)^{-1} \rho\Omega \\
\lambda^\pi &= (I_r - \rho\Omega)^{-1} \rho C H (I_r - \rho H)^{-1}
\end{aligned}$$

Then, we can easily compute the residuals:

$$\begin{aligned}
u_{t+1} &= Z_{t+1} - \mathbb{E}_t Z_{t+1} \\
e_1' u_{t+1} &= r_{t+1} - \mathbb{E}_t (r_{t+1})
\end{aligned}$$

and the news about future cash flows can be obtained as:

$$N_{CF,t+1} = e_1' u_{t+1} + N_{DR,t+1}$$

Note that given a sequence of probabilities or a draw for the MS states and a set of parameters, it is easy and computationally efficient to compute the entire sequences $v^{q,T}$, $v^{\pi,T}$, and u^T:

$$\begin{aligned}
N_{DR}^T &= e_1' w \left[\lambda^q v^{q,T} + \lambda^\pi v^{\pi,T}\right] \\
N_{CF}^T &= e_1' w \left[(I_r + \lambda^q) v^{q,T} + \lambda^\pi v^{\pi,T}\right] \\
u^T &= e_1' w v^{q,T}
\end{aligned}$$

F Recursive Estimates

In the benchmark results presented above, I have used filtered probabilities to pin down agents' beliefs about entering the Great Depression regime and to compute the news. These probabilities would represent the real time probabilities for an agent that has knowledge of all parameters of the model, but not of the regime in place. The choice of endowing the agent with the best possible estimates for the model parameters is consistent with the idea that in reality, agents have more information than the econometrician. Therefore, in using the whole sample, the econometrician is trying to obtain the most accurate estimates of what agents in fact know. It goes without saying that using the whole sample also improves the precision of the estimates.

An alternative approach would consist of assuming that the agents act as econometricians themselves, recursively estimating the MS-VAR as more data become available. I then conduct the following exercise. First, the MS-VAR is estimated over the sample December 1928-January 1965. The corresponding filtered probabilities and parameter estimates are used to compute the news over this initial subsample. Then a month at a time is added, with the result that the subsample keeps expanding until the whole sample is covered. For each of the expanded subsamples, the

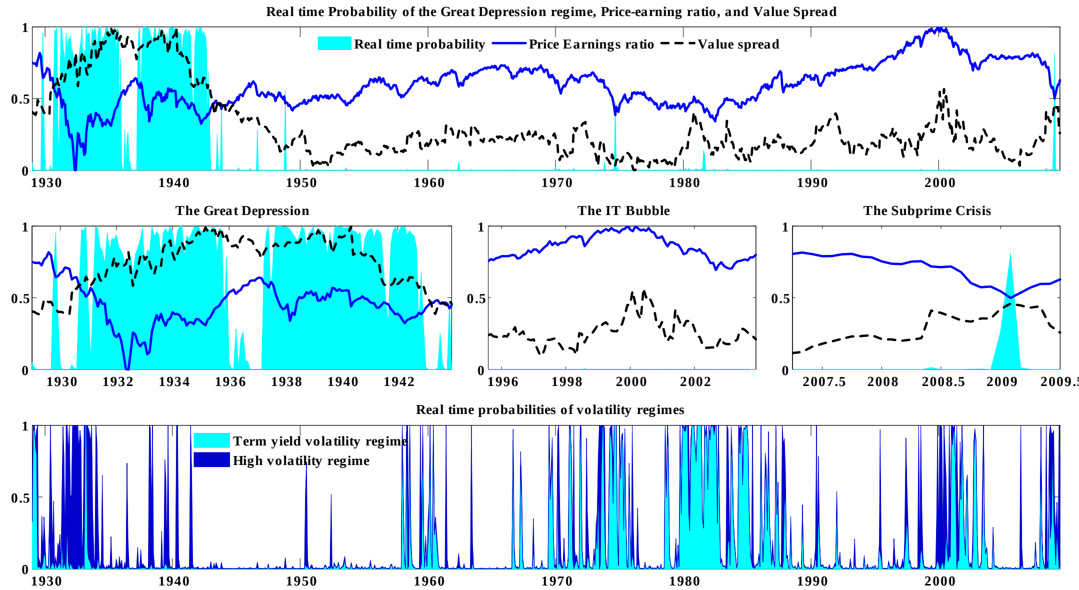

Figure 17: **Real time probabilities.** The figure reports the probabilities for the 1930 regime, the High volatility regime, and the Term yield volatility regime computed in real time starting with an initial sample spanning the period from December 1928 to January 1965. Then a month is added, the model is re-estimated, and the regime probabilities for that month are stored.

model is re-estimated, the *last* value for the regime probabilities and the corresponding parameter estimates are saved, and the news for this additional observation are computed and stored to reflect the updated estimates.[7]

Figure 17 reports the regime probabilities computed in real time for the Great Depression regime, the High volatility regime, and the Term Yield volatility regime. Even in this case, to facilitate the interpretation of the results, the periods corresponding to the Great Depression, the IT bubble, and the Great Recession are enlarged. Note that the results are similar to what is obtained when using the whole sample. In fact, the spike in the probability of the Great Depression regime at the beginning of 2009 is now even larger. In February 2009, the probability of the Great Depression regime computed in real time was 81.03%. This result reinforces the case in favor of the idea that agents might have feared a return to the Great Depression.

With respect to the filtered probabilities obtained using the whole sample, the only noticeable difference consists of an increase in the probability of the Great Depression regime in September 1974. When using the whole sample, the probability of the Great Depression regime during this month is 17.91%, while when using the recursive estimates this probability increases to reach 46.73%. Even in this case, the results can be rationalized in light of historical events. This is a period of time characterized by substantial uncertainty, induced by the end of the first oil shock, the resignation of President Nixon in August 1974 following the Watergate scandal, and the terrorist attack on the TWA Flight 841 from Tel Aviv to New York City after an intermediate

[7] Accordingly, the priors are always set by only using the data available at each point in time.

Sample	R^2	MPE	Risk Aversion
Pre-1964	49.67%	0.7585	3.7665
Post-1964	45.24%	0.6226	14.1706
Whole sample	69.13%	0.3598	9.1992

Table 4: Explanatory power of the ICAPM over different subsamples based on news computed in real time. The explanatory power of the ICAPM is assessed over three different samples: Pre-1964, post-1964, and the whole sample. The table reports R^2, mean pricing error, and the estimate for the coefficient of relative risk aversion. The news are computed by using recursive estimates of the MS-VAR.

stop in Athens.

Table 4 reports the results for the explanatory power of the ICAPM using the news computed in real time. Notice that the R^2 is still very large on both subsamples, even if somewhat lower than when computing the news using all the available information. However, the performance of the model over the whole sample is improved, with an R^2 close to 70%.

G Data for other countries

Value Spread: For each country (UK, Italy, German, Spain and Japan), we obtain from Datastream the following information of each firm, including the delisting ones: Name, Datastream Code, Major Flag, Stock Type, Geography Group, Bourse MNEMONIC, Bourse Name, Unadjusted Price, Number of Shares, Price, Market Value, Price to Book Value. The sampling period is from January 1990 to December 2017 and sampling frequency is monthly.

Below we use UK as an example to illustrate data cleaning steps:
(1) Drop all non-equity constituents, i.e. Stock Type should be "EQ".
(2) Drop all non-major constituents and keep only major listings, i.e. Major Flag should be "Y".
(3) Drop all non-domestic stocks and keep only domestic listings, i.e. Geography Group should be "1" for UK.
(4) Drop all stocks not listed on countries' major exchange(s), i.e. Bourse MNEMONIC should be "LON" and Bourse Name should be "London" for UK.
(5) Drop all the stocks for which the company name contains any suspicious words indicating that the listing may not belong to equities
(6) Drop all the observations of the company when at least one month's unadjusted price is less the £1.00. The threshold £1.00 is arbitrarily chosen and is used to eliminate very small stocks.
(7) Drop all monthly observations of one firm from the end of the sample period back to the occurrence of the first non-zero return. This step is used to deal with delisting firms in Datastream properly.
(8) When there are no observations of the market value (MV) or the two methods of calculating MV yield different results, the MV is replaced with the value calculated by multiplying the unad-

justed price by number of shares.
(9) Drop all monthly negative or infinite book to market values of any firm.

We are now ready to explain the calculation of small stock value spread (VS).

Since June 1991, we sort all stocks of a country into two groups according to size at the end of each June of year t and three groups according to B/M at the end of each December of year t-1. Small stocks (S) are those in the bottom 50% of the June market value, and big stocks (B) are those in the top 50% of the June market value. The B/M breakpoints are the 30th and 70th percentiles based on big stocks (i.e., top 50% of the market value), and are used to divide stocks into growth (G, bottom 30%), neutral (N, middle 40%), and value (V, top 30%). As a result, the independent 2x3 sorts on the size and B/M produce the six portfolios, SG, SN, SV, BG, BN, and BV. We then calculate the monthly value-weighted B/M ratio of SV and SG, which are denoted as $(B/M)_{SV}$ and $(B/M)_{SG}$. The small stock value spread is a monthly time series which is defined as $\log(B/M)_{SV}$-$\log(B/M)_{SG}$.

Term spread: 10-year treasury constant maturity rate and 3-month treasury bill yield of each country come from Fred. Both are in **percent units** and have already been **annualized**. The starting date is January 1990 and the end date is December 2017. The frequency is monthly. All calculations are conducted within Excel.

Log-excess return: Each country's representative monthly index (FTSE 100 for UK, IBEX 35 for Spain, FTSE MIB for Italy, DAX for Germany and Nikkei 225 for Japan) is used to calculate log return. The index data of each country come from Global Financial Data and Yahoo Finance. The priority is given to Yahoo Finance. Global Financial Data is used only when the corresponding series is not available at Yahoo Finance. The 3-month treasury bill yield of each country comes from FRED. The log-excess return is defined as the difference between log-return and the **de-annualized** 3-month treasury bill yield. The starting date of each country's series depends on the availability of each country's data and can be easily found in the Excel file. The end date is December 2017. The frequency is monthly. All calculations are conducted within Excel.

Shiller's log PE ratio: We rely on Datastream Global equity indices. For each country's market, Datastream Global equity indices provide a representative sample of stocks covering a minimum 75 - 80% of total market capitalization, which enables market indices to be calculated. We first calculate the equity index's total earning series (E) of each country using equity index's market value (MV) divided by equity index's Price-Earnings ratio ($PE = \frac{MV}{E}$ by definition in Datastream). Shiller's log PE10 ratio is then defined as the log of the ratio between the equity index's MV and a 10-year moving average of its total earning series (E). The starting date depends on the availability of each country's data and can be easily found in the Excel file. The end date is December 2017. The frequency is monthly. All calculations are conducted within Excel.

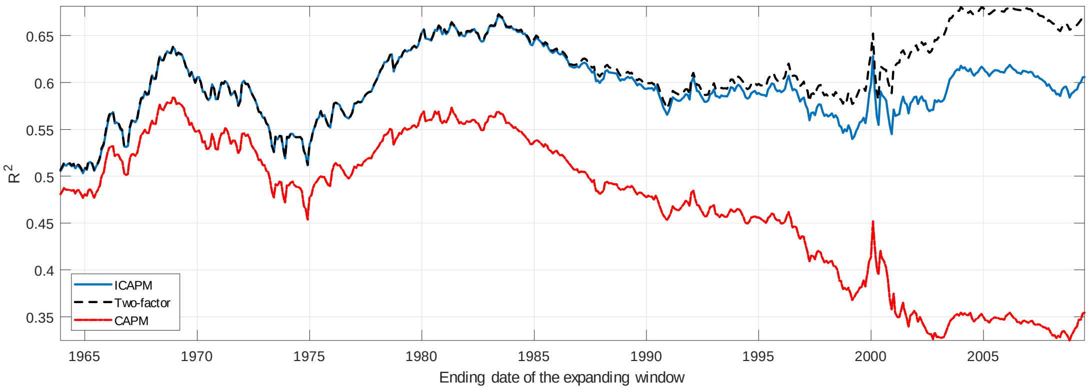

Figure 18: **Explanatory power of the ICAPM over expanding windows starting from the beginning of the sample.** The figure reports the explanatory power as measured by the R^2 of three models: The unrestricted two-factor model, the Intertemporal CAPM, and the traditional CAPM. The betas and average returns are computed over expanding windows that always include the Great Depression. The horizontal axis reports the ending date of the expanding window. The dependent variables are the average returns of the 25 Fama-French portfolios.

H Additional Results

Figure 18 reports the evolution of R^2 for the ICAPM, CAPM, and Two-factor model over expanding windows that always include the Great Depression. For example, 1975 corresponds to the sample January 1929-January 1975. Figure 19 reports the evolution of R^2 for the ICAPM, CAPM, and Two-factor model over expanding windows that always include the Great Recession. For example, 1975 corresponds to the sample January 1975-June 2009. For each subsample, the betas are computed according to the formulas reported in (10) in the paper. The dependent variables are the average returns of the 25 Fama-French portfolios over the same time period. I drop the extreme small-growth portfolio that is often found to be an outlier in asset-pricing models. The results confirm what presented in the paper. As long as the Great Depression and the Great Recession are included in the sample, the ICAPM is able to account for the cross section of asset returns. Instead, the CAPM works well only over the first half of the sample. To see this, note that in 19 the R^2 is extremely up to the point in which the starting date of the sample is the early 1930s.

To understand why modeling the possibility of the Great Depression regime helps in improving the fit of the ICAPM, it is useful to study the behavior of the cash-flow and discount-rate betas in the two cases.[8] The left and right panels of Figure 20 report the cash-flow and discount-rate betas

[8] Campbell and Vuolteenaho (2004) also allow for a lag in the formulas used to compute the betas to control for the possibility that not all stocks in the test-asset portfolios were traded frequently and synchronously. See page 1258 of their paper. In that case, the formulas for the betas become: $\widehat{\beta}_{i,CF} = \frac{\widehat{cov}(r_{i,t},N_{CF,t})}{\widehat{var}(N_{CF,t}-N_{DR,t})} + \frac{\widehat{cov}(r_{i,t},N_{CF,t-1})}{\widehat{var}(N_{CF,t}-N_{DR,t})}$ and $\widehat{\beta}_{i,DR} = \frac{\widehat{cov}(r_{i,t},-N_{DR,t})}{\widehat{var}(N_{CF,t}-N_{DR,t})} + \frac{\widehat{cov}(r_{i,t},-N_{DR,t-1})}{\widehat{var}(N_{CF,t}-N_{DR,t})}$. The results presented in the paper are very similar across the two specifications. I decided to present results based on a beta without a lag because these formulas are more common. The results for the alternative specification are available upon request.

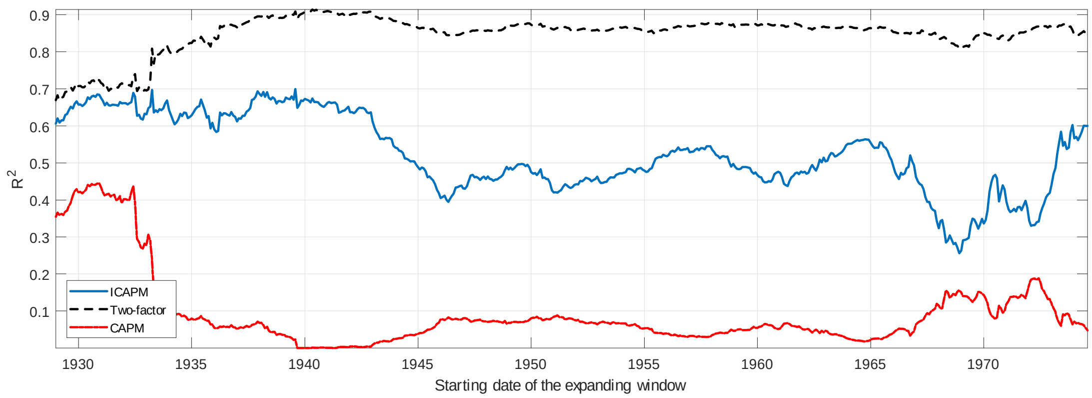

Figure 19: **Explanatory power of the ICAPM over expanding windows starting from the end of the sample.** The figure reports the explanatory power as measured by the R^2 of three models: The unrestricted two-factor model, the Intertemporal CAPM, and the traditional CAPM. The betas and average returns are computed over expanding windows that always include the Great Recession. The horizontal axis reports the starting date of the expanding window. The dependent variables are the average returns of the 25 Fama-French portfolios.

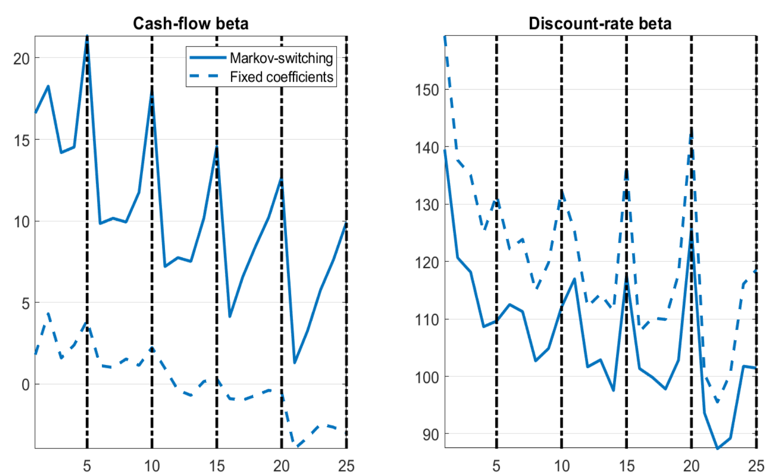

Figure 20: **Betas from MS-VAR and FC-VAR.** The left and right panels report the cash-flow and discount-rate betas computed based on the MS-VAR (solid line) and fixed coefficients VAR (dashed line), respectively. The vertical bars mark the size (from small to large), whereas in each group the five portfolios are organized from growth (low book-to-market ratio) to value (high book-to-market ratio).

Sample	R^2	MPE	Risk Aversion
Pre-1964	50.54%	0.7454	4.8965
Post-1964	54.73%	0.5147	30.8862
Whole sample	60.58%	0.4593	11.3846

Table 5: Explanatory power of the ICAPM over different subsamples. The explanatory power of the ICAPM is assessed over three different samples: Pre-1964, post-1964, and the whole sample. The table reports R^2, mean pricing error, and the estimate for the coefficient of relative risk aversion.

computed based on the MS-VAR (solid line) and fixed coefficients VAR (dashed line), respectively. The vertical bars mark the size (from small to large), whereas in each group the five portfolios are organized from growth (low book-to-market ratio) to value (high book-to-market ratio). The cash-flow betas computed based on the MS-VAR are generally larger and present substantially more variation across the different portfolios, especially across the growth-value dimension (i.e., controlling for size). In turn, this variation helps in accounting for the cross-sectional variation of asset returns.

Summarizing, the analysis of the cross section of asset returns confirms the presence of similarities between the Great Depression and the Great Recession. The latter turned out to be, at least to date, a much less dramatic event. Nevertheless, it seems that both events are key to understanding the cross section of asset returns. To further corroborate this result, Table 5 breaks the sample into two parts, pre and post 1964. Notice that the R^2 is similar across the two subsamples and is quite high. Similarly, the R^2 computed over the whole sample is also very high. Therefore, the results suggest that as long as exceptional events are properly taken into account the ICAPM is able to correctly price the cross section of asset returns. Over the first subsample, the Great Depression plays a key role. Over the second subsample the Great Recession is enough to account for the behavior of the assets during rare events.