BANK OF ENGLAND

Staff Working Paper No. 552

Volatility contagion: new evidence from market pricing of volatility risk

Marek Raczko

September 2015

BANK OF ENGLAND

Staff Working Paper No. 552

Volatility contagion: new evidence from market pricing of volatility risk

Marek Raczko[(1)]

Abstract

This paper proposes a novel approach to assessing volatility contagion across equity markets. I decompose the variance risk premia of three major stock indices into: crash and non-crash risk components and analyse their cross-market correlations. I find that crash-risk premia exhibit higher correlations than non-crash risk premia, implying the existence of volatility contagion. This suggests that investors believe that equity returns will be more highly correlated across countries during market crashes than during more normal times. The main result of the analysis holds when I apply other measures of co-movement as well as when I allow correlation to be time varying. Moreover I document that crash-premia constitute a large portion of the overall variance risk premia, highlighting the importance of crash-risks. Unlike the existing literature, my approach to testing the existence of volatility contagion does not rely on short periods of financial distress, but allows for crash-risk premia to be computed in tranquil times.

Key words: Financial contagion, variance risk premium, tail-risk, equity co-movement, volatility co-movement.

JEL classification: C58, F36, G12, G13, G15.

---

(1) Bank of England. Email: marek.raczko@bankofengland.co.uk

The views expressed in this paper are those of the author, and not necessarily those of the Bank of England or its committees. This article is part of my PhD dissertation at the European University Institute. I am greatly indebted to my supervisors: Evi Pappa and Peter R Hansen. I would also like to thank Pawel Baranowski, Helia Costa, Reinhard Ellwanger, Kristin Forbes, André Gama, Iryna Kaminska, Karen Mayhew, Andrew Meldrum, Jon Relleen, Tomasz Wieladek, Chris Young and other colleagues at the Bank of England. Moreover I am very grateful to participants of the Econometrics working group and the Measuring and Modeling Financial Risk with High Frequency Data 2014 conference at the European University Institute in Florence, the 68th European Meeting of the Econometric Society (ESEM) 2014, Bank of England - LSE workshop on Global Factors and Risks in Financial Markets 2015 and Eighth Annual SoFiE Conference 2015 in Aarhus.

Information on the Bank's working paper series can be found at
www.bankofengland.co.uk/research/Pages/workingpapers/default.aspx

Publications Team, Bank of England, Threadneedle Street, London, EC2R 8AH
Telephone +44 (0)20 7601 4030 Fax +44 (0)20 7601 3298 email publications@bankofengland.co.uk

ISSN 1749-9135 (on-line)

1 Introduction

The recent financial crisis highlighted the high degree of co-movement between international stock markets during crisis periods. This paper studies this co-movement by decomposing international 'variance risk premia' – i.e. the difference between expected market volatility and the volatility implied by equity options (for example in the case of S&P500, the VIX is the implied volatility index) – into two components: one capturing compensation for crash risk and another capturing compensation for 'non-crash' risk. More precisely, I define market crash risk as the risk of an event where the market jumps by at least -10% within one trading day and non-crash risk is defined as any market moves which are not considered to be a market crash. The analysis shows that crash risk premia exhibit higher correlations internationally than non-crash risk premia. This suggests that investors believe that equity returns will be more highly correlated across countries during market crashes than during more normal times.

This paper therefore contributes to the literature on asset price 'contagion' across countries, which - following Forbes and Rigobon (2002) – is defined as an increase in cross-market correlation[1] during times of crisis. While a number of papers have found evidence of this form of contagion (e.g. for equity returns, King and Wadhwani (1990) and Longin and Solnik (1995); for realized equity volatilities, Diebold and Yilmaz (2009); Cipollini et al. (2013); and for option-implied equity volatilities, Cipollini et al. (2013), other studies, after correcting for estimator biases (e.g. Forbes and Rigobon (2002), Longin and Solnik (2001), and Corsetti et al. (2005)) find no evidence of contagion. Dungey and Zhumabekova (2001) point out that the primary difficulty is that periods of turmoil are usually short and consequently span only a small portion of the observed sample. Moreover the choice of dates for the financial turmoil 'regime' might also lead to inconsistent or inefficient estimates.

The novel approach developed in this paper avoids many of the drawbacks associated with distinguishing changes in correlation during short crisis periods. This is due to the fact that I look directly at market pricing of crash risk, which can be computed during tranquil or crisis period. More precisely, I decompose variance risk premia[2] into components compensating for crash and

[1]Traditionally correlation of stock market indices or asset prices were analyzed, but in this study I focus on the co-movement of volatilities of major stock market indices.

[2]Variance Risk Premium is the premium that markets require for the risk of a change of uncertainty. This premium is calculated as a difference between the statistical measure of market volatility (empirically measured by the realized volatility) and the risk neutral implied

non-crash states in the United States, the United Kingdom and euro-area, by applying a modified version of the method of Bollerslev and Todorov (2011b) to the S&P500, FTSE100 and Eurostoxx50, respectively. This allows me to compare the co-movement of premia that compensate for crash events with the co-movement of premia for the remainder of the variance risk (i.e. 'non-crash risk premia').[3] I find that crash risk premia exhibit higher cross-country correlation than non-crash risk premia. This suggests that investors believe that the correlation of equity returns will be higher in tail events than in more normal times, which provides strong evidence for the market contagion hypothesis.

Moreover, crash-risk premia correlations are elevated, relative to the correlation of non-crash risk premia, even when I account for time-varying correlation using the Dynamic Conditional Correlation model of Engle (2002). Hence the main result of the paper is robust to possible time-variation in the strength of international relationships. In fact, cross-country correlations of crash-risk premia are time-varying, yet they remain quite stable over time. I find that even though individual market crash risk premia are very sensitive to adverse market events (e.g. Russian default, LTCM collapse, Lehman Brothers bankruptcy, Sovereign default crisis, etc.), their international co-movement remains relatively stable.

Aside from providing important evidence for market contagion in times of crisis, the high correlation of tail risk premia has important implications for both financial market practitioners and policymakers. First, it shows that the potential gains from portfolio diversification are smaller than would be expected when not accounting for tail-dependency, as cross-country hedging will not be effective during times of crisis. Models that do not capture this feature seem likely to overestimate the gains from international diversification and the degree of investors' home bias.

Second, policymakers are likely to be particularly concerned with the impact of domestic monetary policy on perceptions of crash-risks. Hattori et al. (2015) studied the impact of US quantitative easing (QE) on crash risk perceptions, finding that QE resulted in a statistically significant decrease in crash premia. My analysis shows that policy that reduces crash-risk premia is likely to have a global impact. This implies that US QE might have large spillover effects on other equity markets and consequently on other economies through its impact

volatility (empirically measured by the options implied volatility index, ex. VIX).

[3]Bollerslev et al. (2012) or Londono (2014) show that the Variance Risk Premia are dominated by a global component, yet they do not look into the split of the VRP into the tail- and non-tail risk related premia.

on reducing global crash risk premia. The analysis developed in this study suggests that an interesting direction for future research is to investigate this particular global aspect of QE.

The remainder of the paper is organized as follows. Section 2 briefly describes the method and Section 3 characterizes the dataset used for the analysis. Section 4 describes the results and Section 5 concludes.

2 Methodology

The methodology in this study comprises of three parts. First, I define the concept of Variance Risk Premium (VRP) and I show how it is measured using daily data on options and 5-minute frequency intra-day data on index futures prices. Second, I describe how to decompose VRP into the part related to crash risk and the part related to non-crash risk, using techniques developed by Bollerslev and Todorov (2011a). Given that my S&P500 options data differ from theirs (in that my option dataset exhibits longer average maturities) and that I am also extending their calculations to new datasets, namely FTSE100 and Eurostoxx50, I also describe my modification of the original methodology. Finally I describe the co-movement measures used in the study. Specifically, I use the Dynamic Conditional Correlation model of Engle (2002) to analyse potentially time-varying correlations between premia across equity markets.

2.1 Variance Risk Premium (VRP)

Many financial studies have shown that not only equity returns, but also volatilities (risks) of those returns are time-varying. This basic fact of non-constant volatility means that this is an additional source of investment risk. The Variance Risk Premium (VRP) is the compensation that market requires for this additional risk. In fact, financial markets have already developed tools to hedge the risk of volatility increase. VRP can be traded using variance swaps (see Demeterfi et al. (1999) for details). These instruments simply swap future unknown realized variance for current option implied variance.

In technical terms, the VRP is measured as the difference between the physical expectation (the P-measure) of the realized quadratic variation of returns and the risk-neutral expectation (the Q-measure) of the quadratic variation of returns.

$$VRP_t = \frac{1}{T-t}\left(E_t^P(QV_{[t,T]}) - E_t^Q(QV_{[t,T]})\right) \tag{1}$$

The physical expectation (the P-measure) of the quadratic variation is simply the statistical $T-t$ periods ahead forecast. Quadratic variation under P is measured as the realized variance (RV) based on 5-minute frequency intra-day prices of index futures.[4] This approach has been strongly advocated by Sheppard et al. (2013), who showed that this is the best variance estimator. Moreover in this study, following Bollerslev et al. (2009), I use simple naive expectations of the realized variance as a proxy for the forecast of realized variance. This approach should be effective as variance exhibits large persistence, exemplified by volatility clustering.[5]

$$E_t^P(QV_{[t,T]}) = \sum_{i=t-(T-t)}^{t} RV_i \tag{2}$$

Risk-neutral expectations of the quadratic variation (the Q-measure) are measured using daily data on the panel of options. Those data enable us to calculate the model-free option-implied variance of future prices. This type of variance measure reflects the expected variance implied by option prices under the assumption of risk neutral market pricing. In more technical terms this measure is derived under the assumption that the stochastic discount factor is constant and equal to the inverse of the risk-free interest rate. This means that the Q-measure of the variance combines investors' expectations of the future variance with their risk preferences (see Figlewski (2012)).[6] The most classical

[4] In order to adjust for the overnight price changes daily realized variance is re-scaled by the constant proportion of overnight change.

[5] More recently, however Bekaert and Hoerova (2014) or Kaminska and Roberts-Sklar (2015) show that the naive forecast can be improved if the forecasting method models separately the continuous and the jump part of the volatility. Furthermore the forecast might be improved even more by the use of option implied volatility data. Yet, given that the focus of this study is the decomposition and cross correlation of VRPa, it seems that simple naïve expectations forecast would work well.

[6] Simple coin flipping game might be a great example to understand the difference between Q- and P- measure of the probability distributions. Say, the game pays EUR 100 in case the flip yields heads and 0 in the other case. The P–measure would correspond to the actual distribution, hence both events have probabilities equal to 0.5. In order to determine the Q-measure of probabilities we need to know the price of the game. Say, an economic agent is willing to pay EUR 30 for that game. Under the assumption of risk-neutrality this would mean that the distribution of the probability should be 0.3 for heads and 0.7 for tails. The difference between those two measures of probabilities simply reflects agents risk aversion.

example of a model-free Q-measure of volatility is the VIX index.[7]

My Q-measure of the quadratic variation only slightly differs from the VIX index.[8] Both measures use approximation to calculate implied volatility for a fixed time horizon. Yet, unlike the VIX which uses only two different option maturities to calculate approximated values, I use the whole available set of different option maturities. Moreover, in contrast to the VIX methodology which approximates linearly quadratic volatility, I approximate option prices using Carr and Wu (2003) polynomial and based on theoretical option prices I calculate the implied volatility.[9] This change in the calculation method is motivated by two factors. First, the set of data used in this study, suffers from a small number of very close to maturity options, hence the VIX methodology would imply linear extrapolation from the two options with quite distant maturities. This seems inappropriate, especially when dealing with options capturing large jump probabilities. Second, I wanted to keep my measure consistent with the decomposition of the VRP presented in subsection 2.2.1.

Equation 3 describes the formula for the Q-measure of the quadratic variation, once the theoretical 14-day to maturity options are calculated:

$$E_t^Q(QV_{[t,T]}) = \frac{2}{T-t}\sum_i \frac{\Delta K_i}{K_i^2} e^{(T-t)r} Q(K_i) - \frac{1}{T-t}\left[\frac{F}{K_0} - 1\right]^2 \tag{3}$$

In my calculations an option's time to maturity $T - t$ is fixed to 14 days (it is always quoted as a fraction of a year). The forward index level F is calculated based on the index level at a given moment and the respective (14 day) risk-free interest rate r. K_0 denotes the first strike price below the forward index level F of the panel of options. K_i is the strike price of ith out-of-the-money option; a call if $K_i > K_0$ and a put if $K_i < K_0$; both put and call if $K_i = K_0$. ΔK_i is simply a mid-point between two strike prices: K_{i-1} and K_{i+1}. The price of the option $Q(K_i)$ for a given strike price is either a price of the call option $C(K_i)$ if $K_i > K_0$ or a price of a put option $P(K_i)$ if $K_i < K_0$. The entire equation 3 is exactly the same as the one used to calculate the VIX index (see Chicago Board Options Exchange White Paper (2009)).

Finally, as shown in equation 1, VRP is measured as the difference between

[7] To obtain implied variance, VIX index has to be divided by 100 and squared.

[8] In fact the correlation of my measures with volatility indices: VIX, VFTSE and VStoxx is very high and amounts roughly to 95%.

[9] Please refer to the Appendix A for more details on the approximation.

the two expectations, hence it reflects investors' attitude towards risk – the so called risk appetite. The decomposition of this risk enables us to understand what drives the VRP: crash-events or more "normal" type of equity return movements. In the next section I describe the basic assumptions needed to calculate how much of the VRP is attributed to market crash risk.

2.2 Tail-premia measures

The Bollerslev and Todorov (2011a) methodology, which is applied in this paper, requires that the underlying asset price follows a very general jump-diffusion process.[10] It implies that the asset price dynamics (in case of this study price of futures for the underlying index F_t) follows a stochastic differential equation:

$$\frac{dF_t}{F_t} = \alpha_t dt + \sigma_t dW_t + \int_R (e^x - 1)\tilde{\mu}(dt, dx) \tag{4}$$

where α_t denotes the drift, σ_t denotes the instantaneous volatility and W_t is a standard Brownian motion. The first two elements of the sum depict the continuous part of the dynamics. The third part of the sum describes jumps or discontinuities of the asset price dynamics, where the $\tilde{\mu}(dt, dx)$ is the so-called compensated jump measure. The jump part may for example follow a Poisson process as in the Merton (1976) model. But, in the case of this study, there is no need to limit ourselves to any parametric distribution - neither for the continuous, nor for the jump part. In fact, for our analysis, the most important feature of this model is the additive separability of the continuous and the jump components.

Both the diffusion and the jump part of the asset price dynamics will have their parallels in the process describing asset price variance. Consider the quadratic variation of the logs of asset prices over the [t,T] time interval:

$$QV_{[t,T]} = \int_t^T \sigma_s^2 ds + \int_t^T \int_R x^2 \mu(ds, dx) \tag{5}$$

where the first component $\int_t^T \sigma_s^2 ds$ is the volatility of the continuous process

[10] This type of process is very common in the financial literature, mainly due to the fact that it fits the actual data very well. Moreover, it allows prices to exhibit discontinuous patterns, which in turn, justifies the existence of markets for financial options in theoretical finance models (for some discussion of merits of jump-diffusion models please refer to Tankov and Voltchkova (2009)).

and the second component $\int_t^T \int_R x^2 \mu(ds, dx)$ denotes the volatility generated by the discontinuous part. In principle the first part should be responsible for the volatility generated by the "smaller" (continuous) movements in the asset prices, whereas the second part would depict volatility generated by the "larger" asset price movements (jumps).

Quadratic variation equation 5 implies that the VRP, defined by equation 1, will simply be a sum of two differences: the difference between P and Q expectations of the continuous part of the quadratic variation and the difference between P and Q expectations of the jump part of the quadratic variation:

$$\begin{aligned} VRP_t = & \frac{1}{T-t}\left(E_t^P(\int_t^T \sigma_s^2 ds) - E_t^Q(\int_t^T \sigma_s^2 ds)\right) \\ & + \frac{1}{T-t}\left(E_t^P \int_t^T \int_R x^2 \mu(ds, dx)) - E_t^Q(\int_t^T \int_R x^2 \mu(ds, dx))\right) \end{aligned} \tag{6}$$

I need all the above presented structure to define the variance risk premium solely attributed to the market crash risk – $VRP(\tilde{k})$. This measure describes the contribution to the respective P- and Q- measures of quadratic volatility by asset price drops higher than a certain threshold k. In my study I define market crash as a state when asset prices fall by at least 10%. This implies that my threshold level $k = ln(0.9)$ and consequently $\tilde{k} = 0.9$. The price change of 10% can definitely be considered as a large move, hence it will only be reflected by the discontinuous part of the VRP. Consequently my VRP(k) measure depends only on the jump parts:

$$\begin{aligned} VRP_t(\tilde{k}) = & \frac{1}{T-t}\left(E_t^P\left(\int_t^T \int_{x<k} x^2 v_s^P(dx)ds\right)\right) \\ & - \frac{1}{T-t}\left(E_t^Q\left(\int_t^T \int_{x<k} x^2 v_s^Q(dx)ds\right)\right) \end{aligned} \tag{7}$$

Finally on the basis of the VRP(k) and the total VRP, I can also define a truncated volatility measure VRP(tr). This measure will capture the part of the variance risk premium that is attributed to the remaining non-crash risk:

$$VRP_t(tr) = VRP_t - VRP_t(0.9) \tag{8}$$

Having defined tail-risk premia, the next two sub-sections briefly describe how to calculate Q- and P- measures from the data.

2.2.1 Risk-Neutral (Q) Measures

The most difficult part of the Q-measure estimation is to pin down the process of the time-varying jump density $v_t^Q(dx)$. In order to construct a time-varying measure with as few assumptions regarding its structure as possible, I estimate it non-parametrically from the options data. Therefore I assume the following for jump density:

$$v_t^Q(dx) = (\varphi_t^- 1_{\{x<0\}})v^Q(x)dx \tag{9}$$

where φ_t^- denotes an unspecified stochastic process of temporal variation of the jump arrivals and $v^Q(x)$ is an unspecified time-invariant density. Yet, the methodology of Bollerslev and Todorov (2011b) allows us to estimate tail-volatilities $E_t^Q\left(\int_t^T \int_{x<k} x^2 v_s^Q(dx)ds\right)$ even under those very general assumptions. First of all they calculate model-free risk neutral measures from the panel of options data. Second, using the Extreme Value Theory (EVT) those measures are used to estimate Generalized Pareto Distribution (GPD) parameters (namely: scale (σ) and shape (ξ) parameters) and the average jump intensities $E(\frac{1}{T-t}E_t^Q(\int_t^T \varphi_s ds))$ through a just identified GMM system. This allows us to fully describe the time invariant part of the jump intensity $v^Q(x)$ for large price changes. Third, using fixed parameters for the GPD, the time varying jump intensities are backed out to fulfill exactly the moment conditions. Finally, using the estimated parameters the Q-measure of the tail-volatility is calculated for a given threshold k.

I describe the risk neutral jump-tail measures in detail as here I deviate slightly from the original Bollerslev and Todorov (2011b) framework. They propose a model-free risk-neutral jump tail measure:

$$LT_t^Q(k) = \frac{e^r P_t(K)}{(T-t)F_t} \tag{10}$$

where $k = ln(\frac{K}{F})$ is the log-moneyness, $P_t(K)$ is the price of a put option, K is the option strike price and F_t is the price of the underlying futures. This measure captures solely the jump risk as long as two conditions are fulfilled. First the options have to be deeply out of the money. Bollerslev and Todorov (2011b) use moneyness levels of $\{0.9000\ 0.9125\ 0.9250\}$, which should guarantee enough distance from the underlying to capture only the jump risk. Second the option needs to be close to maturity. Bollerslev and Todorov (2011b) use options that have median of 14 days to maturity. In my calculations I follow the same levels of option moneyness, but the dataset used in this study has much longer median maturity of options (see Table 8 in Appendix A). This means that my model-free risk-neutral jump tail measures might be "contaminated" by the diffusive part of the process. In fact, when I applied the exact Bollerslev and Todorov (2011b) methodology, my jump tail measure for S&P500 was substantially larger when the options had longer maturities relative to the original study.

In order to circumvent this problem I use a panel of options with different maturities for a given moneyness level to fit the polynomial describing the time-decay plot of option price. Carr and Wu (2003) show that this polynomial should approximate the time-decay of options no matter whether the underlying process contains jumps or not. This approximation allows me to calculate the theoretical price of the 14-days-to-maturity option. Appendix A provides details on the approximation method as well as some robustness checks.

Once I have the theoretical 14-days-to-maturity option price, I construct the same risk-neutral jump tail measure. In this case the pattern of my jump tail measure closely resembles the original one of Bollerslev and Todorov (2011b).

Generalized Pareto Distribution (GPD) parameters are estimated using the simple non-linear GMM procedure of Hansen and Singleton (1982). The exact moment conditions are described in the Appendix B. The basic principle is that for left tail I have 3 parameters to estimate and jump-tail measures for 3 different levels of moneyness, hence the system is just identified.

2.2.2 Objective (P) Measures

Analogous to the Q-measure estimation, the key issue in estimation of the P-measure is to pin down the time-varying jump density v_t^P. Unfortunately it is not possible to estimate the intensity fully non-parametrically, simply because I do not have three different points of the curve on the same day. Consequently I assume an affine model of the jump intensity. Following Bollerslev and Todorov (2011a) I assume that the temporal variation of the volatility is a function of the stochastic volatility σ_t^2 of the continuous part:

$$v_t^P(dx) = (\alpha_0^- 1_{\{x<0\}} + \alpha_1^- 1_{\{x<0\}} \sigma_t^2) v^P(x) dx \tag{11}$$

This implies that I have to estimate four parameters that are constant across time (namely: scale (σ) and shape (ξ) parameters of the GPD that characterizes

$v^P(x)$, and α_0 and α_1). Moreover I have to get the estimate of the time-varying stochastic volatility σ_t^2. Here again, I follow closely Bollerslev and Todorov (2011a) framework.

First I estimate continuous volatility using Mancini (2001) idea of truncated volatility. All intra-day asset price movements below a certain threshold contribute to the continuous volatility whereas the ones above the threshold contribute to the jump volatility. The truncation threshold is time-varying to capture the effects of the volatility clustering. The threshold is a function of the past continuous volatility. Moreover the daily pattern of volatility is also taken into account. For each index I estimate the average volatility for a given time. On that basis I calculate the time of the day volatility multiplier that either increases or decreases the threshold. For more details on the realized volatility calculations please refer to the Appendix C.

Second I select a threshold level, which is always higher than the maximum threshold used to determine continuous volatility. I select a threshold of 0.6% for all the indices. On the basis of this threshold I can mark observations that are definitely jumps in the whole sample. Then I use estimated continuous volatility along with matrices indicating jumps (the ones determined by 0.6% threshold) to estimate all four parameters in question. Again the estimation is done using the GMM framework (for details on the exact moments specification please refer to the Appendix D).

Finally, once all the parameters are calculated I calculate the tail-volatilities $E_t^P\left(\int_t^T \int_{x<k} x^2 v_s^P(dx)ds\right)$ for the threshold of $ln(0.9)$ to match the tail-volatilities for the Q-measure.

2.3 Co-movement measures

The main result of this analysis is based on the measures of co-movement of $VRP(0.9)$ as well as $VRP(tr)$ across equity markets (i.e. three indices: S&P500, FTSE100 and Eurostoxx50). In order to keep the analysis simple and yet powerful the main result is based on the simple r-Pearson correlation coefficient. The main finding is based on comparing unconditional correlations of crash risk premia to unconditional correlations of non-crash risk premia.

The correlation coefficient is known to be sensitive to outliers however, which is why I also report two non-parametric measures of co-movement: Kendall's τ and Sperman's ρ. Those measures are used as a robustness check of the main finding.

Market correlations are renowned to be time varying, hence as a final robustness check to my main correlation matrix I allow correlation to be time-varying. In order to capture a more complex dynamic correlation structure, I apply the Dynamic Conditional Correlation (DCC) model of Engle (2002). This model helps me not only to overcome the problem of time-varying correlation, but to control for the heterogeneity of individual shocks. The model looks at the conditional correlations of innovations, enabling me to gauge how shocks co-move across markets and is given below:

$$y_t = C + \sum_{k=1}^{K} A_k y_{t-k} + \epsilon_t \tag{12}$$

$$E_{t-1}(\epsilon_t \epsilon_t') = \Sigma_t \tag{13}$$

$$\Sigma_t = D_t R_t D_t \tag{14}$$

$$R_t = diag(Q_t)^{-1/2} Q_t diag(Q_t)^{-1/2} \tag{15}$$

$$Q_t = (1 - \lambda_1 - \lambda_2)\bar{R} + \lambda_1 \tilde{\epsilon}_{t-1} \tilde{\epsilon}_{t-1}' + \lambda_2 Q_{t-1} \tag{16}$$

$$\bar{R} = E[\tilde{\epsilon_t} \tilde{\epsilon_t}'] \tag{17}$$

$$\tilde{\epsilon}_t = D_t^{-1} \epsilon_t \tag{18}$$

The DCC model requires the level equation to be parsimonious, hence in the benchmark case I use VAR(SIC) processes to describe variables' levels (see equation 12), where the number of lags is selected on the basis of Schwartz information criteria.[11] The vector of variables in equation 12 contains either all three crash-risk VRP(0.9) or all three non-crash-risk VRP(tr).

The conditional covariance matrix (equation 13) is decomposed into the matrix of individual conditional standard deviations D_t and conditional correlation matrix R_t (see equation 14). Conditional standard deviation matrix D_t is a di-

[11] As an additional robustness check I have also used other models, namely: AR(1), AR(SIC) and VAR(1), but this changes did not yield qualitatively different results.

agonal matrix where each element on the diagonal simply represents a square root of individual variances which are modelled as the GARCH(1,1) process. Transformation of the conditional correlation matrix (see equation 15) guarantees that the matrix has ones on the diagonal. Quasi conditional correlation (see equation 16) is a weighted average of the unconditional sample correlation $\bar{R}$ (see equation 17) and the previous period cross product of 'corrected' innovations (see equation 18) and the previous period conditional quasi correlation. The specification of the equation 16 nests the Constant Conditional Correlation (CCC) model of Bollerslev (1990), hence allowing for direct testing of the time varying correlation assumption. Should λ_1 and λ_2 parameters were jointly statistically insignificant, then the correlation between innovations would be constant over time.

3 Data

The dataset used in this study allows me to replicate the US results of Bollerslev and Todorov (2011b) as well as to extend their calculations to the UK and euro-area. Accordingly, US calculations are based on the S&P500 index, UK on the FTSE100 index and euro-area on the Eurostoxx50 index. The Q-measure (implied distribution) is based on a daily panel of options, whereas the P-measure (statistical measure) is based on intra-day (5 minutes) data on traded futures, obtained from Thomson Reuters. Finally, correlation calculations are conducted on the weekly averages, as the daily data contained too much noise.

3.1 Options

I use options data collected by the Bank of England from Chicago Mercantile Exchange (CME), Eurex Exchange and London International Financial Futures and Options Exchange (LIFFE) for S&P500, Eurostoxx50 and FTSE100 index options, respectively. The data are sampled with a daily frequency. The data for S&P500 and FTSE100 options span January 1995 to December 2013. Unfortunately the data span for the Eurostoxx50 is shorter and covers only January 1999 to December 2013. This sample still allows me to cover major period of market turmoil (for US and UK only: LTCM, Russian and Asian crises and for all three markets: dotcom bubble burst, accounting scams and the great recession period for all indices).

I apply a standard set of filters to the options data before any calculations take place. The set of filters is based on programmes used by the Bank of

England which are in line with the ones used in Carr and Wu (2009).

3.2 Intra-day data of index futures

I use the intra-day data provided by Thomson Reuters. The data are sampled at a 5-minute frequency. This frequency allows me to capture price jumps limiting the impact of the microstructural noise. In fact Sheppard et al. (2013) show that realized variance based on 5-minute frequency data is the best estimator of the realized variance across different assets.

For S&P500 and FTSE100 I use the data spanning January 1996 to December 2012, whereas for Eurostoxx50 the data only spans January 1999 to December 2012. The range of the dataset for the S&P500 is unfortunately shorter than in the Bollerslev and Todorov (2011b) paper, hence the parameter estimates might differ. In terms of trading time during each day, for each series I have tried to pick a time period for which I had data throughout all the dates. Consequently, my time windows are: for S&P500 - 81 observations (from 8.30 to 15.10), for FTSE100 - 94 observations (from 8.15 to 16.00) and for Eurostoxx50 - 81 observations (from 9.15 to 15.55).

4 Results

Before I go to the main result of the paper, i.e. the analysis of the co-movement of risk premia across equity markets, I would like to describe briefly the estimates of the GPD parameters under Q and P probability measures.

4.1 Q-measure

Table 1 summarizes parameter estimates for the risk-neutral Q-measure. Parameters are precisely estimated, as can be seen from standard errors. The first two parameter estimates describe the time-invariant parameters of the GPD, ξ - the shape parameter and σ - the scale parameter. The larger those parameters are the thicker the tails of the distributions.

It is clear that, according to my estimates, the tail of the FTSE100 index distribution is the thinnest, as all parameters are the smallest from all three indices. In case of S&P500 and Eurostoxx50 the results are more ambiguous. The scale parameter is marginally bigger for the Eurostoxx50, but the shape parameter is bigger for the S&P500. This means that, even though for smaller values Eurostoxx50 tail is thicker, for larger jumps the S&P500 tail is thicker.

The estimates of the average jump intensity parameters αv, calculated at

-7.5% price jumps, are only comparable between S&P500 and FTSE100, as Eurostoxx50 estimates were calculated on the different sample. Yet, again those estimates show that FTSE100 options exhibit the smallest tail-risk.

It is easier to interpret annualized average jump intensities presented in Table 2, as they swiftly summarize the impact of all three parameters on the tails. For example, the results in Table 2 read that we should expect about 3 jumps of -10% in four years for S&P500 index. All those numbers indicate that those probabilities are higher than the actual, even extreme, price changes observed on the futures markets. Actually, a -10% index jump has not been observed in any of the analysed samples. This is likely to be a manifestation of the fact that risk premia are embedded in the Q distribution.

Moreover it is also very interesting to note that the crash contribution (index jump of at least -10%) to the overall Q measure of variance is 41%, 33% and 46% for S&P500, FTSE100 and Eurostoxx50, respectively. Of course those numbers are averages specific to the analysed samples.

One could also enquire how those estimates for the left tail compare to those of the right tail. Analogue calculations for the right tail can be found in the Appendix F. It is worth noting that under Q-measure tail distributions are highly skewed to the left, manifesting the so called volatility 'smile'.

Table 1: Q-measures estimation results

	S&P500	Eurostoxx50	FTSE100
ξ	0.2744	0.2693	0.2313
	(0.0092)	(0.0096)	(0.0094)
σ	0.0563	0.0590	0.0527
	(0.0007)	(0.0007)	(0.0006)
αv	1.2425	1.4751	1.1012
	(0.0152)	(0.0208)	(0.0138)

Notes: Table reports estimated parameters of the generalized Pareto distribution under the risk neutral Q-measure: ξ is the estimate of the shape parameter and σ is the estimate of the scale parameter. αv is the estimate of the average annualized jump intensity of -7.5% jump in the price level. The estimates are based on S&P500 and FTSE100 options data from January 1996 to December 2013 and Eurostoxx50 options data from March 2002 to December 2013. The log-moneyness of options used to estimate parameters were 0.9000, 0.9125 and 0.9250. Estimates standard errors are reported in parentheses.

Table 2: Q-measure: annualized jump intensity estimates

Jump Size	S&P500	Eurostoxx50	FTSE100
$<$-7.5%	1.2425	1.4751	1.1012
$<$-10%	0.7554	0.9153	0.6445
$<$-20%	0.1393	0.1764	0.0999

Notes: Table reports annualized average jump intensities under the Q measure i.e. implied by option prices. Jump sizes are in terms of percentage changes in price levels. In the case of S&P500 and FTSE100 averages are calculated from January 1996 to December 2013, and for Eurostoxx50 averages are calculated from March 2002 to December 2013. All the reported figures are based on generalized Pareto distribution estimates reported in Table 1.

4.2 P-measure

Table 3 summarizes estimation results for the objective, 'physical', P-measure. The first two parameters describe the time-invariant shape (ξ) and scale (σ) of the GPD, similarly as in case of the Q-measure. Unfortunately those estimates are not directly comparable with the ones from the Q-measure, as they were calculated at a different thresholds.

Estimates of those two parameters do not differ substantially across analysed markets. In contrast to estimates of the Q-measue, under the P-measure the FTSE100 tail seems to be the thickest. This might be partially explained by the fact that this market is considered to be the least liquid of the three.

The α_0 and α_1 parameters describe the affine process driving jump-intensities under the P-measure. The significance of the estimates of α_1 parameters for all three markets indicate that jump-intensities are in fact time varying and closely connected to the actual continuous volatility.

As in the case of Q-measure, it is worth looking at the average jump intensities for the P-measure. Table 4 shows that a single day -10% market crash is an extremely rare event. In the case of the FTSE100 index, for which the P-measure is the most leptokurtic, estimated annualized jump intensities imply that we would only observe 1 such crash in 100 years. This is even more striking when compared to roughly 55 such events in 100 years implied by the Q-distribution. This yet again underscores the impact of the risk-aversion on the Q-measure.

Moreover, the contribution of market crash to the total variance under the P-measure is much smaller that under the Q-measure and amounts to 0.05%, 0.01% and 0.11% for S&P500, Eurostoxx50 and FTSE100, respectively. This implies that the compensation for crash events is larger than that for the 'regular' volatility.

Finally, Appendix G contains analogous results for the right tail of the distribution. Interestingly I find that the tails under P-measure are also skewed to the left, but much less than under the Q-measure. This contrasts with the Bollerslev and Todorov (2011a) findings, which note a skew towards the right tail. It can be explained by the fact that the sample I use also covers the period of the great recession.

Table 3: P-measure estimation results

	S&P500	Eurostoxx50	FTSE100
ξ	0.2500	0.2305	0.2596
	(0.0766)	(0.0701)	(0.0406)
100σ	0.1594	0.1819	0.1624
	(0.0155)	(0.0164)	(0.0083)
α_0	-0.0016	-0.0016	-0.0025
	(0.0001)	(0.0001)	(0.0001)
α_1	0.0329	0.0346	0.0406
	(0.0005)	(0.0006)	(0.0007)

Notes: Table reports estimated parameters of the generalized Pareto distribution under the physical P-measure: ξ is the estimate of the shape parameter and σ is the estimate of the scale parameter. α_0 and α_1 are estimates of the parameters in equation 11 which links jump intensities to the time-varying continuous volatilities. The estimates are based on high-frequency 5-minute futures prices from January 1996 to December 2012 for S&P500 and FTSE100, and from January 1999 to December 2012 for Eurostoxx50. Estimates standard errors are reported in parentheses.

Table 4: P-measure: annualized jump intensity estimates

Jump Size	S&P500	Eurostoxx50	FTSE100
$<$-7.5%	0.0069	0.0082	0.0343
$<$-10%	0.0020	0.0022	0.0102
$<$-20%	0.0001	0.0001	0.0004

Notes: Table reports annualized average jump intensities under the P measure i.e. based on the high frequency data estimation. Jump sizes are in terms of percentage changes in price levels. In the case of S&P500 and FTSE100 averages are calculated from January 1996 to December 2012, and for Eurostoxx50 averages are calculated from January 1999 to December 2012. All the reported figures are based on generalized Pareto distribution estimates reported in Table 3.

4.3 Variance Risk Premia and Crash Risk Premia

All the observed Variance Risk Premia (VRP) are on average negative (see Table 5). This is due to the fact that option implied variances (Q-measures) are on average larger than realized variances (P-measures). Moreover VRP are also volatile and persistent. These results show that markets are charging significant and time-varying premia for the risk of future changes of the variance of the asset prices.

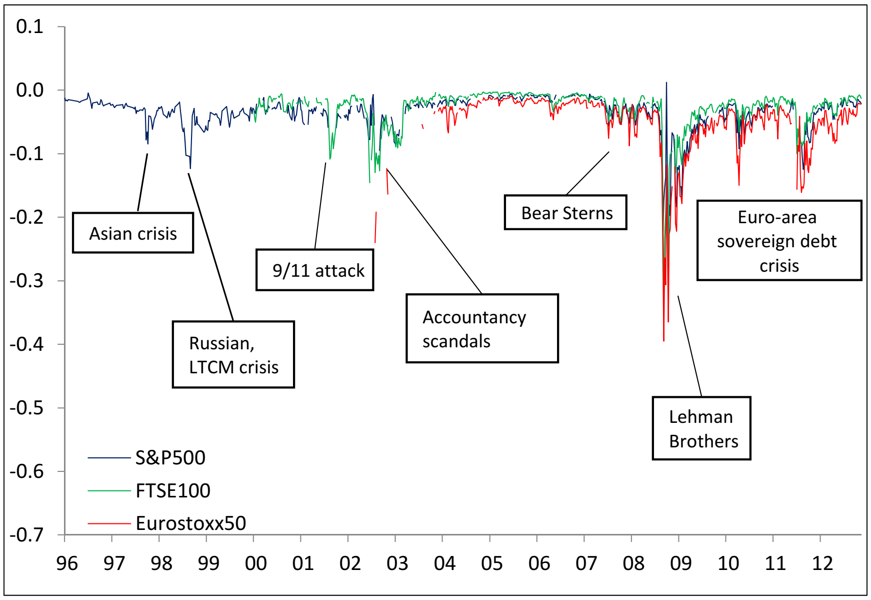
VRP seem to be quite closely co-moving across those three equity markets, hinting that the premium might be globally driven. Premia magnitudes are also very sensitive to major market events, such as accountancy scandals, Bear Sterns melt down, Lehman Brothers' bankruptcy or sovereign default crisis (see Figure 1).

Table 5: Summary statistics for Variance Risk Premia

	S&P500	Eurostoxx50	FTSE100
	VRP		
Mean	-0.0368	-0.0537	-0.0281
Median	-0.0250	-0.0390	-0.0185
Std dev.	0.0373	0.0511	0.0366
	VRP(0.9)		
Mean	-0.0336	-0.0433	-0.0249
Median	-0.0132	-0.0237	-0.0114
Std dev.	0.0579	0.0622	0.0463
	VRP(tr)		
Mean	-0.0032	-0.0104	-0.0032
Median	-0.0087	-0.0136	-0.0060
Std dev.	0.0296	0.0186	0.0139

Notes: Table reports summary statistics for Variance Risk Premia (VRP), crash-risk VRP(0.9) and non-crash-risk VRP(tr). VRP is defined as the difference between the statistical expectations (P-measure) of variance and option implied (Q-measure) variance, calculated on the basis of high frequency 5-minute futures prices data and daily option prices data, respectively. On average VRP is negative, indicating that on average implied variance is higher than the statistically expected variance, showing that market participants are risk averse. Crash-risk VRP is the part of the premia that is required solely to hedge market crash risk, defined here as a -10% jump in the underlying index. VRP(tr) is the residual premium that is required for non-crash risk. Calculations are based on weekly averages from March 2002 to December 2012.

Figure 1: Variance Risk Premia



Notes: The Figure shows evolution of Variance Risk Premia (VRP) over time for S&P500, Eurostoxx50 and FTSE100. Labels depict major global market events. Missing data for S&P500 and Eurostoxx50 are due to gaps in the options datasets. The figure represents weekly averages of VRP.

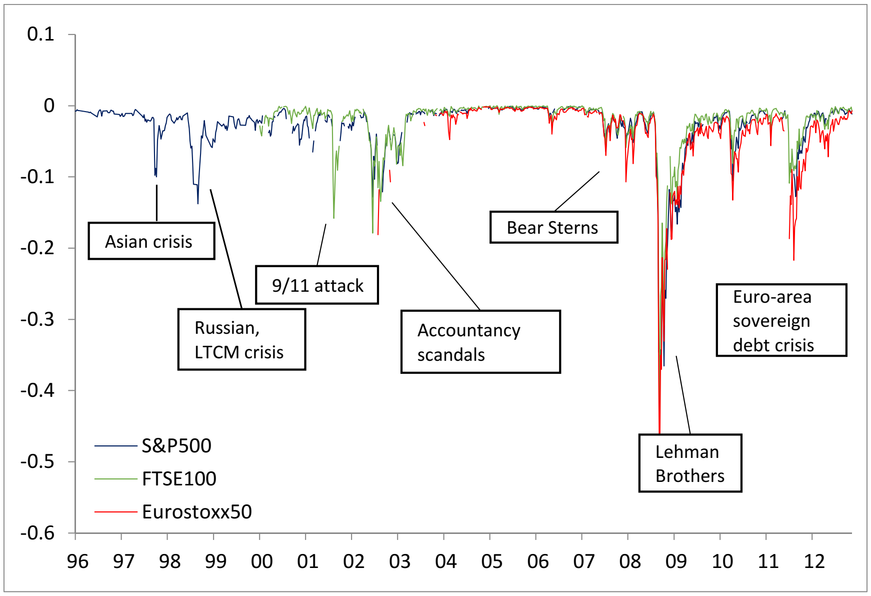
$VRP(0.9)$ attributed solely to the market crash seem to exhibit the same features as the total VRP. They are also negative, persistent and volatile. They also react sharply to major market events. Actually one may easily note that VRP and crash $VRP(0.9)$ are co-moving for all three indices. This is not surprising as crash $VRP(0.9)$ constitute large fractions of the total VRP.

In fact, on average, it captures 91%, 81% and 89% of VRP for S&P500, Eurostoxx50 and FTSE100, respectively. Those results are in line with the study of Bollerslev and Todorov (2011b) who found that 88% of the S&P500 VRP is driven by the crash premium. It should be also noted that these results are driven by premia values during market turmoil times, as ratios of median crash $VRP(0.9)$ to median total VRP are much smaller, though still substantial. More precisely, they amount to 53%, 61% and 62% for S&P500, Eurostoxx50 and FTSE100, respectively. Yet both sets of numbers clearly indicate high importance of crash premia in the total risk compensation.

The result of high impact of crash risk on the market compensation for risk is in line with rare disasters literature. Rietz (1988) and subsequently Barro

(2006), ? and Wachter (2013) highlight this phenomenon in theoretical macro-financial models. From that perspective my finding simply empirically reinforces their analysis.

Figure 2: Crash-Risk Variance Risk Premia (0.9)



Notes: The Figure shows evolution of crash-risk VRP (0.9) (i.e. the premium for holding volatility risk associated with a -10% jump in the price of the underlying index futures) for S&P500, Eurostoxx50 and FTSE100. Labels depict major global market events. Missing data for S&P500 and Eurostoxx50 are due to gaps in the options datasets. The figure represents weekly averages of VRP (0.9).

4.4 New evidence on contagion

The main question of this paper is the existence of the volatility contagion. This question is answered by the comparison of cross-market correlations of crash risk premia VRP(0.9) against correlations of the non-crash risk premia VRP(tr).

Table 6 summarizes the key result of this paper – **crash risk premia co-move by more than the premia for non-crash risk across all three equity indices.** This indicates that large negative events (market crashes) have more global impact than other 'regular' events. This proves the existence of volatility contagion on equity markets. It should be once more underlined that, in contrast to the existing literature, this test for market contagion does not depend on a crash event, but is based on market pricing of crash risk.

Table 6: Pairwise correlations of VRP(0.9) and VRP(tr)

	Pearson's correlation	
	VRP(tr)	VRP(0.9)
S&P500 Eurostoxx50	0.5214	0.9559
S&P500 FTSE100	0.3026	0.9631
FTSE100 Eurostoxx50	0.2508	0.9624

Notes: Table reports pairwise correlations of three index pairs for two measures: crash-risk VRP(0.9) (i.e. the premium for holding volatility risk associated with a -10% jump in the price of the underlying index futures) and non-crash-risk VRP(tr) (i.e. the premium for holding volatility risk not related to the market crash). Pairwise correlation are calculated on a common sample of weekly data for all three indices from March 2002 to December 2012. The table clearly shows that correlations of crash-risk premia VRP(0.9) are substantially higher than correlations of the non-crash risk premia VRP(tr).

Moreover, crash premia VRP(0.9) seem to be driven by a common factor. In fact, simple principal component analysis indicates that the first principal component of three crash premia VRP(0.9) describes 97% of total data variability, whereas in case of the reminder of the volatility premia VRP(tr) it amounts to 87%.

Crash premia VRP(0.9) are quite volatile and susceptible to market adverse events (like the collapse of Lehman Brothers), hence one might suspect that the high correlation results are driven solely by outliers. In order to check whether presented results are robust to outliers, I also look at two non-parametric measures of dependence, namely Kendall's τ and Spearman's ρ. Table 7 shows that even under those measures of dependence, crash premia VRP(0.9) are more closely co-moving than the non-crash premia VRP(tr). This reinforces the existence of volatility contagion.

Kendal's τ for crash premia are markedly lower than in the case of Pearson's correlations, but still higher than the correlations of non-crash premia, whereas Sperman's ρ dependence measures are in line with Pearson's correlation numbers. If anything, this simple robustness exercise indicates that the outliers are rather decreasing the non-crash premia correlations, but still they are always lower than the correlation of crash premia.

Table 7: Non-parametric dependence measures of VRP(0.9) and VRP(tr)

	Kendall's τ		Spearman's ρ	
	VRP(tr)	VRP(0.9)	VRP(tr)	VRP(0.9)
S&P500 Eurostoxx50	0.5648	0.7972	0.7154	0.9430
S&P500 FTSE100	0.3919	0.7996	0.5334	0.9475
FTSE100 Eurostoxx50	0.3088	0.7888	0.4120	0.9369

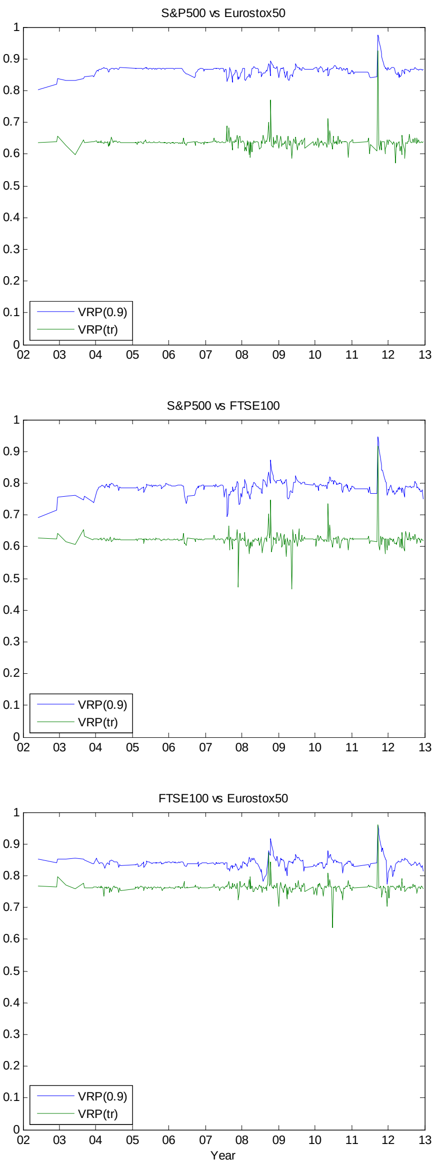
Notes: Table reports non-parametric pairwise dependence measures of three index pairs for two premia measures: crash-risk VRP(0.9) (i.e. the premium for holding volatility risk associated with a -10% jump in the price of the underlying index futures) and non-crash-risk VRP(tr) (i.e. the premium for holding volatility risk not related to the market crash). Two non-parametric measures are Kendall's τ and Sperman's ρ. These measures are used as they should be more robust to outliers than the simple correlation coefficient. Dependence measures are calculated on a common sample of weekly data for all three indices from March 2002 to December 2012. The table confirms findings documented by a simple correlation coefficient in Table 6.

Market correlations are perceived to be unstable over longer periods of time. In order to overcome this problem, as the last robustness check, I have extended my analysis to allow for the dynamic correlations. A simple rolling window analysis, presented in Appendix H, shows that correlations in question are indeed unstable. Moreover, Forbes and Rigobon (2002) pointed out that this type of simple analysis might be biased due to heterogeneity of individual shocks.

As mentioned earlier, I solve both issues by looking at the VAR(SIC)-DCC(1,1) model, as it allows for time varying: correlations and individual volatilities. The number of lags for each series is determined by the Schwartz information criteria. The model shows that the conditional correlations of dynamic innovations are indeed time-varying, for crash risk premia VRP(0.9) as well as for non-crash premia VRP(tr). In fact, the data rejected the Constant Conditional Correlation model of Bollerslev et al. (1988) that is embedded in the DCC specification.

Even though correlations are time-varying the main result of the paper remains in place as correlations of crash premia VRP(0.9) are always higher than correlations of non-crash premia VRP(tr) (see Figure 3). Finally, the result of higher co-movement of crash premia also holds when the level equation follows different processes, namely AR(1), AR(SIC), VAR(1).[12] This highlights the

Figure 3: Time varying conditional correlations calculated on the basis of weekly data by VAR(SIC)-DCC(1,1) model.



Notes: Figures show the dynamic correlations of three index pairs for two premia measures over time: crash-risk VRP(0.9) (i.e. the premium for holding volatility risk associated with a -10% jump in the price of the underlying index futures) and non-crash-risk VRP(tr) (i.e. the premium for holding volatility risk not related to the market crash). Dynamic correlations are calculated using the Dynamic Conditional Correlation model of Engle (2002). The model is based on a common sample of weekly data for all three indices from March 2002 to December 2012. The level equation of either VRP(0.9) or VRP(tr) for all three indices is modelled jointly as a VAR process, where the number of lags is selected using Bayesian information criterion.

robustness of the key result.

5 Conclusions

In this study I showed that the volatility premia investors require to compensate for crash risks are more closely co-moving across different equity markets than volatility premia required for non-crash risks. This result implies that investors perceive crash risks to have more global impact than other risks, hence pointing to market contagion. This study uses a novel approach to assess the volatility contagion. Unlike previous studies that compare market co-movement in crisis times with co-movement in 'tranquil' times, I compare co-movement of market variance premia for market crash risk with co-movement of non-crash risk variance premia. This allows me to circumvent many of the econometric issues that existing studies suffer from. More precisely, I do not have problems with dating crisis periods or having short crisis data samples. Finally, it should be underlined again that the main result of the paper is robust to different measures of premia co-movement as well as to possible time variation in correlations.

[12] Graphs of dynamic correlations under different level equations, can be found in Appendix I

References

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

Bekaert, G. and M. Hoerova (2014). The vix, the variance premium and stock market volatility. *Journal of Econometrics 183*(2), 181–192.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *Review of Economics and Statistics*, 498–505.

Bollerslev, T., R. F. Engle, and J. M. Wooldridge (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, 116–131.

Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected stock returns and variance risk premia. *Review of Financial Studies 22*(11), 4463–4492.

Bollerslev, T. and V. Todorov (2011a). Estimation of jump tails. *Econometrica 79*(6), 1727–1783.

Bollerslev, T. and V. Todorov (2011b). Tails, fears, and risk premia. *Journal of Finance 66*(6), 2165–2211.

Bollerslev, T., V. Todorov, and S. Z. Li (2012). Jump tails, extreme dependencies, and the distribution of stock returns. *Journal of Econometrics*.

Carr, P. and L. Wu (2003). What type of process underlies options? a simple robust test. *Journal of Finance 58*(6), 2581–2610.

Carr, P. and L. Wu (2009). Variance risk premiums. *Review of Financial Studies 22*(3), 1311–1341.

Cipollini, A., I. L. Cascio, and S. Muzzioli (2013). Volatility co-movements: a time scale decomposition analysis. Technical report, Universita di Modena e Reggio Emilia, Facoltà di Economia" Marco Biagi".

Corsetti, G., M. Pericoli, and M. Sbracia (2005). Some contagion, some interdependence: More pitfalls in tests of financial contagion. *Journal of International Money and Finance 24*(8), 1177–1199.

Demeterfi, K., E. Derman, M. Kamal, and J. Zou (1999). More than you ever wanted to know about volatility swaps. *Goldman Sachs quantitative strategies research notes 41*.

Diebold, F. X. and K. Yilmaz (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *Economic Journal 119*(534), 158–171.

Dungey, M. and D. Zhumabekova (2001). Testing for contagion using correlations: some words of caution. *Pacific Basin Working Paper Series* (2001-09).

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics 20*(3), 339–350.

Figlewski, S. (2012). What is risk neutral volatility? *Available at SSRN 2183969*.

Forbes, K. J. and R. Rigobon (2002). No contagion, only interdependence: measuring stock market comovements. *Journal of Finance 57*(5), 2223–2261.

Hansen, L. P. and K. J. Singleton (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica*, 1269–1286.

Hattori, M., A. Schrimpf, and V. Sushko (2015). The response of tail risk perceptions to unconventional monetary policy. *Available at SSRN 2566769*.

Kaminska, I. and M. Roberts-Sklar (Forthcoming, 2015). Global factor in variance risk premia and local bond pricing. *Bank of England Staff Working Paper*.

King, M. A. and S. Wadhwani (1990). Transmission of volatility between stock markets. *Review of Financial Studies 3*(1), 5–33.

Londono, J. M. (2014). The variance risk premium around the world. *Available at SSRN 2517020*.

Longin, F. and B. Solnik (1995). Is the correlation in international equity returns constant: 1960–1990? *Journal of International Money and Finance 14*(1), 3–26.

Longin, F. and B. Solnik (2001). Extreme correlation of international equity markets. *Journal of Finance 56*(2), 649–676.

Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping brownian motion. *Giornale dell'Istituto Italiano degli Attuari 64*, 19–47.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics 3*(1), 125–144.

Rietz, T. A. (1988). The equity risk premium a solution. *Journal of Monetary Economics 22*(1), 117–131.

Sheppard, K., L. Liu, and A. J. Patton (2013). Does anything beat 5-minute RV? a comparison of realized measures across multiple asset classes. *University of Oxford, Department of Economics Economics Series Working Papers* (645).

Tankov, P. and E. Voltchkova (2009). Jump-diffusion models: a practitioner's guide. *Banque et Marchés 99*, 1–24.

Wachter, J. A. (2013). Can time-varying risk of rare disasters explain aggregate stock market volatility? *Journal of Finance 68*(3), 987–1035.

Appendices

Appendix A - Time-decay approximation

The dataset used in this study has one substantial drawback - the time to maturity of options is much longer than in the Bollerslev and Todorov (2011b) study (see Table 8), except for FTSE100. Consequently the estimator of the tail measure could be "contaminated" by the diffusion process. This in turn may bias my estimates of the Generalized Pareto Distribution leading to an inaccurate inference about tail-risk premia. In order to circumvent this problem I use all available maturities of options to estimate the time-decay patterns. This allows me to calculate the theoretical value of options that have 14 days to maturity. I choose this number of days to maturity to match exactly the median number of days to maturity in the Bollerslev and Todorov (2011b) study.

Table 8: Maturities of the closest to maturity options

Index	Minimum	Maximum	Median
BT: S&P500	5	x	14
S&P500	6	75	33
FTSE100	5	29	15
Eurostoxx50	5	74	36

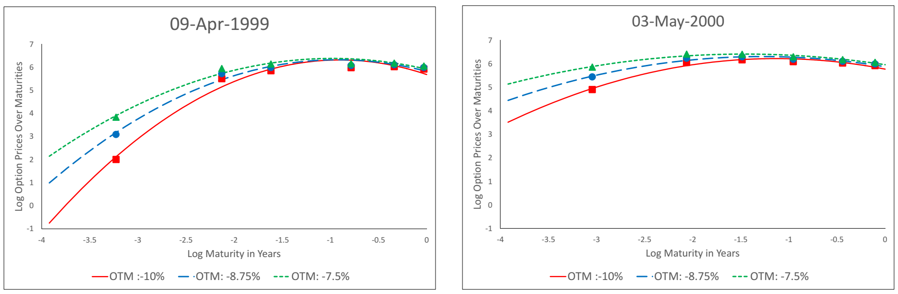
Notes: Table reports minimum, maximum and median days-to-maturity of the closest to maturity option used in my dataset as well as options used in the original Bollerslev and Todorov (2011b) study. Only the median days to maturity for the FTSE100 roughly matches the one of Bollerslev and Todorov (2011b).

Out-of-the-money options at the maturity have zero value. However, the order of convergence over time to that value depends largely on the process governing the underlying asset's price dynamics. Carr and Wu (2003) showed that the time decay (or the order of convergence) of out-of-the money options is dominated by the presence of jumps. They showed that if the price of the underlying asset follows a jump process or a jump-diffusion proces, then the value of the out-of-the-money option will converge more slowly to zero than in the case of a strict diffusion process. They also showed that the time decay of option prices can be closely approximated by the following polynomial:

$$ln\left(\frac{P}{T}\right) = a(lnT)^2 + b(lnT) + c \tag{19}$$

This approximation equation is valid regardless whether the underlying process exhibits jumps or not. If the underlying equity process has no jumps the fitted line should have a greater slope close to the zero maturity (as the price of the option is falling faster than the time to maturity), whereas if it exhibits jumps the time-decay plot should be flatter (see Figure 4). In this study I fit this polynomial for each day of the data - since the perception of the jump probability might change over time. The fitted line allows me to calculate the theoretical option value for the exact 14 days to maturity.

Figure 4: Time-decay plots with fitted polynomial



Notes: Figures depict time-decay plots for FTSE100 index options. Markers represent actual option prices while lines represent the fitted polynomial (see equation 19). The left panel shows time-decay of FTSE100 index options on the 9th of April 1999, a very tranquil period when jumps were very unlikely. The right panel shows time-decay FTSE100 index options in 3rd of May 2000, a more volatile period when jumps were more likely. The dates are chosen to match the graph in Carr and Wu (2003) article so as to make a comparison. In each figure the three lines, from the bottom to the top, represent three moneyness levels of out-of-the-money option prices: -10% (red, solid line), -8.75% (blue, dashed line) and -7.5% (green, dotted line).

The number of options used in the approximation varies over time and is driven by data availability. I use from 4 to 6 option maturities to fit the polynomial - Table 9 shows details for each index.[13] I should expect to get the best results for the FTSE100 index as its option data displays the highest quality - shortest maturities and most of the dataset is covered by 6 maturities. However given that the S&P500 index is the only one present in the original Bollerslev and Todorov (2011b) study I will use this to start my robustness check.

Table 9: The proportion of maturity nodes in the data

Number of options	S&P500	FTSE100	Eurostoxx50
6	17%	91%	86%
5	18%	9%	7%
4	65%	0%	6%

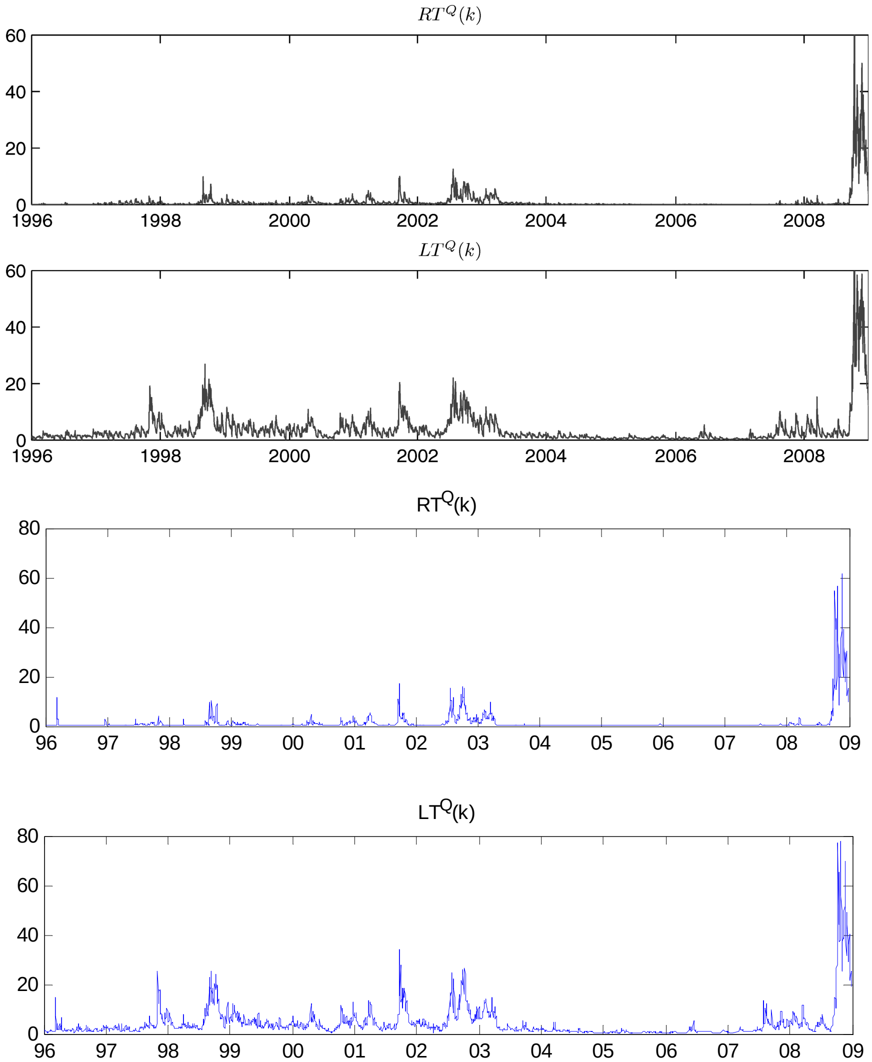
Notes: Table reports the structure of available options data used in this study. For S&P500 and FTSE100 options data ranges from January 1996 to December 2013 with many missing points for S&P500 in the earlier part of the sample. The Eurostoxx50 options data ranges from March 2002 to December 2013, and also exhibits missing data in the earlier part of the sample. Missing data points are due to the fact that for those dates only 3 options with different maturities were available. Those data points were discarded.

First of all it might be noted that the dynamics of tail measures calculated on the basis of the approximation follows nearly the same pattern as the one of Bollerslev and Todorov (2011b) (see Figure 5). The two biggest differences are a jump in the tail measure in early 1996 that is only present in my calculation

[13] It should be noted that for certain periods I only had 3 options at my disposal. Those data-points were removed from the dataset, leading to significant number of missing datapoints for S&P500 and Eurostoxx50 series especially visible before 2003.

and a more pronounced response of my tail measure to the 'dotcom bubble' burst in the late 2001. Unfortunately I do not have the original time-series data of tail measures computed by Bollerslev and Todorov (2011b), so I cannot calculate any goodness-of-fit measure. Yet, I can compare the GMM estimation results (see Table 10). The estimates of the GPD parameters are very close to each other especially for the left tail, as this tail is estimated with a higher accuracy. The only substantial difference is slightly higher estimates of the jump intensity parameters. However as one may note from the final results of the structure of the jump probabilities, the differences are not very large (see Table 11). Judging by the sole comparison of my results to the ones of Bollerslev and Todorov (2011b), it appears that the approximation does a very good job.

Figure 5: Tail measures comparison between Bollerslev and Todorov (2011b) study and this article.



Notes: Figures depict tail measures for the left $LT^Q(k)$ and the right tail $RT^Q(k)$ calculated from S&P500 options, where $k = 0.9$ and $k = 1.1$ for the left and the right tail, respectively. The first two panels are from the Bollerslev and Todorov (2011b) article, where available closest-to-maturity options were used. Last two panels are based on my own calculations, where theoretical 14-day-to-maturity options are used to calculate tail measures.

Table 10: GMM estimates of Q-tail parameters

	BT: S&P500		S&P500	
	LT	RT	LT	RT
ξ	0.2581	0.0793	0.2570	0.0615
	(0.0282)	(0.0147)	(0.0130)	(0.0161)
σ	0.0497	0.0238	0.0513	0.0242
	(0.0021)	(0.0010)	(0.0009)	(0.0006)
αv	0.9888	0.5551	1.1431	0.7266
	(0.0525)	(0.0443)	(0.0142)	(0.0156)

Notes: Table compares parameter estimates obtained by Bollerslev and Todorov (2011b), the BT: S&P500 column, and estimates obtained in this article where I use approximated 14-day to maturity options. LT and RT denote estimates for the left tail and right tail, respectively. Estimates are based on the same sample ranging from January 1996 to June 2007. It should be noted that the sample used by me has some missing data points prior to January 2003, moreover in the main article the calculations are based on the more up-to-date sample. Standard errors are reported in parenthesis.

Table 11: Annualized jump intensities implied by Q-tail distributions

Jump Size	BT: S&P500	S&P500
>7.5%	0.5551	0.7266
>10%	0.2026	0.2666
>20%	0.0069	0.0082
<-7.5%	0.9888	1.1431
<-10%	0.5640	0.6627
<-20%	0.0862	0.1052

Notes: Table compares estimated average jump intensities over January 1996 to June 2007 sample obtained by Bollerslev and Todorov (2011b), the BT: S&P500 column, and estimates obtained in this article where I use approximated 14-day to maturity options. Jump sizes are in terms of percentage changes in price levels. It should be noted that the sample used by me has some missing data points prior to January 2003, moreover in the main article the calculations are based on the more up-to-date sample.

Yet, it is still important to see how well the approximation does with other indices. Here I cannot rely on others results, as to the best of my knowledge I am the first one to estimate these measures for other indices, namely Eurostoxx50 and FTSE100. Consequently I have looked at two fit measures and the volatility of the theoretical prices for different sets of maturity structures (see Table 12 and Figure 6). The simple goodness-of-fit measure (R^2) does not seem to be a good metric. It is exceptionally high for all indices as the dataset has only a small number of nodes. The MAPE of the fit evaluated only at the 14-days to maturity also seems to be very small, except for the FTSE100. In that case the MAPE value is ballooned by having a denominator very close to zero. It is very difficult to drive any conclusions from those simple fit metrics as they are based on an insufficient number of data points for each polynomial.

Table 12: The fit of the time-decay polynomial

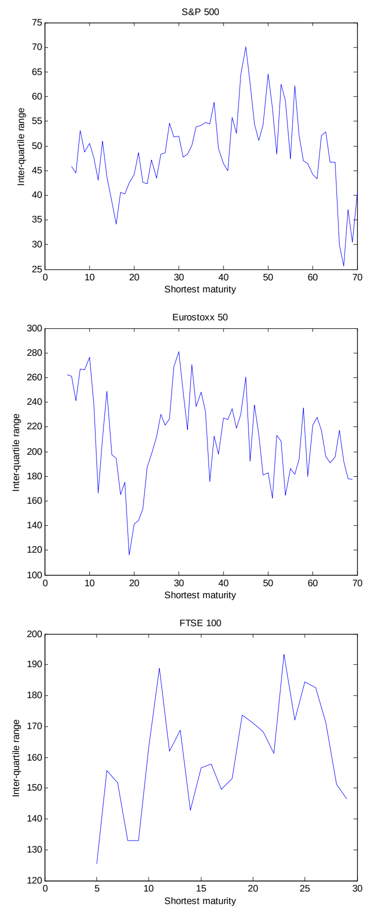
	S&P 500	FTSE 100	Eurostoxx 50
R^2 of the polynomial for 6 different moneyness levels			
Minimum	98.95%	86.04%	99.53%
Average	99.99%	99.91%	99.99%
Percentage error of predicted price for 14-days to maturity option			
MAPE	0.25%	2.86%	0.36%
Maximum	3.47%	56.43%	3.81%

Notes: Table reports different measures of fit of 14-day to maturity option prices by the estimated polynomial. The top part of the table reports minimum and average determination coefficients R^2 of the daily regressions of the option time decay polynomial. The bottom part reports average and maximum observed percentage error of polynomial implied 14-day to maturity option price relative to the actual 14-day to maturity option price. Naturally, the bottom part uses only the observations where the actual 14-days-to-maturity options data were available.

In order to overcome the problem of an insufficient number of data points I have looked at volatilities of theoretical 14-day to maturity option prices approximated using option prices with different maturities. In principle the volatility of the theoretical price should not depend on the set of nodes uses in the approximation (at least not too much). Of course if I extrapolate the 14-day price from a big "distance" the error of fit might generate a higher error than if I use actual maturities very close to the 14 days. Nonetheless it seems informative to investigate how much of the extra volatility is being caused by having distant maturities while performing the approximation. Figure 6 presents inter-quartile ranges for theoretical 14-day prices.[14] The volatility of the theoretical price rises across minimum volatility pointing to certain losses caused by the approximation, but the increase does not seem to be excessive.

[14]Inter-quartile range is being used instead of standard deviations to make the measure robust to outliers.

Figure 6: Inter-quartile ranges of the theoretical 14-day to maturity option prices for different set of maturities used in the approximation



Notes: Figures depict inter-quartile ranges of the approximated (theoretical) 14-day to maturity option price. For all indices -10% out of the money options were approximated and used for the approximation. Horizontal axis denotes the shortest maturity used for the approximation. Inter-quartile ranges are used instead of standard deviations to circumvent the impact of outliers. For S&P500 and FTSE100 data ranged from January 1996 to December 2013 with many missing points for S&P500 in the earlier part of the sample. For Eurostoxx50 data ranged from March 2002 to December 2013, also exhibiting missing data in the earlier part of the sample.

All in all it seems that the approximation is giving a good proxy for the original method especially as the estimates do not differ too much from the original study.

Appendix B - GMM conditions to estimate GPD parameters in the Q measure

The aim of the GMM estimation for the Q-measure is to find the following vector of parameters for each tail:

$$\theta^Q = [\alpha_Q^{\pm}\bar{v}_{\psi}^{Q\pm}(tr^{\pm}); \xi_Q^{\pm}; \sigma_Q^{\pm}]$$

Those parameters are found by fulfilling the following three moment conditions:

$$E(LT_t^Q(k)) = \alpha_Q^{-}\bar{v}_{\psi}^{Q-}(tr^{-})\frac{\xi_Q^{-}}{\xi_Q^{-}+1}\left(e^k\right)^{1+1/\xi_Q^{-}}\left(\frac{\xi_Q^{-}}{\sigma_Q^{-}}\right)^{-1/\xi_Q^{-}} *$$

$$*_2F_1\left(1+\frac{1}{\xi_Q^{-}}; \frac{1}{\xi_Q^{-}}; 2+\frac{1}{\xi_Q^{-}}; \frac{tr^{-}\frac{\xi_Q^{-}}{\sigma_Q^{-}}-1}{e^{-k}\frac{\xi_Q^{-}}{\sigma_Q^{-}}}\right)$$

$$E(RT_t^Q(k)) = \alpha_Q^{+}\bar{v}_{\psi}^{Q+}(tr^{+})\frac{\sigma_Q^{+}}{1-\xi_Q^{+}}\left(1+\frac{\xi_Q^{+}}{\sigma_Q^{+}}(e^k-1-tr^{+})\right)^{1-1/\xi_Q^{+}}$$

where ${}_2F_1$ is a hypergeometric function and $E(LT_t^Q(k))$ and $E(RT_t^Q(k))$ are sample averages of the introduced tail-measures, for right and left tails respectively. Standard errors of estimates are obtained using the delta method.

Parameter estimates for the right tail are presented in the Appendix F.

Appendix C - Realized and continuous variation

In order to compute P-measure components of the VRP and the crash risk VRP(k) we need to compute realized variance (RV) and extract the continuous variation (σ_t^2) from it.

Daily RV is computed using 5-minute high frequency intra-day data on prices of index futures (F_t). More specifically, RV is a sum of squared changes of log prices of index futures (f_t) scaled-up by the average overnight contribution O:

$$RV_t = \left[\sum_{i=1}^{n-1}(f_{t+i\Delta} - f_{t+(i-1)\Delta})^2\right] * O$$

where n is the number of daily prices available in the data, Δ denotes the 5-minute time increment, and the overnight scaling factor O is computed in the following way: $O = 1 + \frac{\sum_{t=1}^{T}(f_t - f_{t-1+(n-1)\Delta})/T}{\sum_{t=1}^{T}(\sum_{i=1}^{n-1}(f_{t+i\Delta} - f_{t+(i-1)\Delta}))/T}$.

In calculating continuous variation (σ_t^2) I follow directly the methodology suggested by Mancini (2001). Essentially the calculations resemble those for RV, with the exception that only the change in log prices that are smaller than the time-varying threshold α_t are added:

$$\sigma_t^2 = \left[\sum_{i=1}^{n-1} (f_{t+i\Delta} - f_{t+(i-1)\Delta})^2 \mathbb{I}_{\{|f_{t+i\Delta} - f_{t+(i-1)\Delta}| \leq \alpha_t\}} \right] * O$$

where $\mathbb{I}$ is an indicator function amounting to one if the absolute change falls below the threshold α_t and zero otherwise.

The time-varying threshold α should take into account the intra-day volatility patterns as well as time varying volatility across days. In order to control for the first one I estimate daily volatility patterns for each futures index. First I set a general truncation level $\bar{\alpha}$ for the whole dataset, so that my calculations are not biased by outliers. The general truncation level $\bar{\alpha}$ is based on the average sample volatility for 5-minute log-price change, measured by the bi-power variation:

$$\bar{\alpha} = 3\sqrt{\frac{\pi}{2}} \sqrt{\frac{1}{T} \sum_{t=1}^{T} \sum_{i=2}^{n-1} |f_{t+i\Delta} - f_{t+(i-1)\Delta}||f_{t+(i-1)\Delta} - f_{t+(i-2)\Delta}|} \left(\frac{1}{n}\right)^{0.49}$$

In turn, this threshold is used to calculate the average log-price variation for every 5-minutes of the trading day (only for the data falling below the threshold):

$$Var_i = \frac{\sum_{t=1}^{T} (f_{t+i\Delta} - f_{t+(i-1)\Delta})^2 \mathbb{I}_{\{|f_{t+i\Delta} - f_{t+(i-1)\Delta}| \leq \bar{\alpha}\}}}{\sum_{t=1}^{T} \mathbb{I}_{\{|f_{t+i\Delta} - f_{t+(i-1)\Delta}| \leq \bar{\alpha}\}}}$$

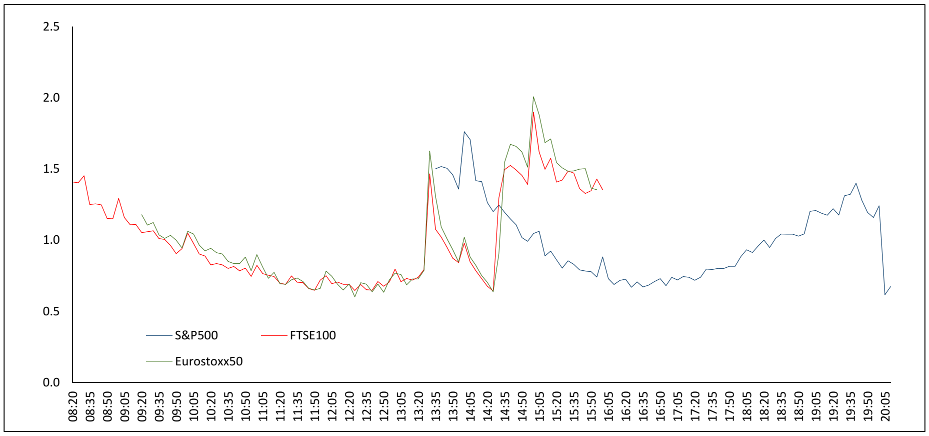
Finally, in order to obtain the time-of-day factor (TOD_i) I normalize each 5-minutes variation (Var_i) by the total sample truncated variation (Var_{TOT}):

$$Var_{TOT} = \frac{\sum_{t=1}^{T} \sum_{i=1}^{n-1} (f_{t+i\Delta} - f_{t+(i-1)\Delta})^2 \mathbb{I}_{\{|f_{t+i\Delta} - f_{t+(i-1)\Delta}| \leq \bar{\alpha}\}}}{\sum_{t=1}^{T} \sum_{i=1}^{n-1} \mathbb{I}_{\{|f_{t+i\Delta} - f_{t+(i-1)\Delta}| \leq \bar{\alpha}\}}}$$

$$TOD_i = \frac{Var_i}{Var_{TOT}}$$

Figure 7 plots TOD factors for all three analyzed indices on the standardized GMT scale. All time of day volatility patterns roughly exhibit a U shape, showing that most of the volatility comes at the beginning and closing of trading time. In addition European indices, Eurostoxx50 and FTSE100, also experience a large increase in volatility at the opening time of the New York Stock Exchange. Whereas the closure of the European trading has a rather minuscule impact on the S&P500 daily volatility pattern.

Figure 7: Time-of-day factor



Notes: The figure shows the estimated time-of-day factor for the S&P500, FTSE100 and Eurostoxx50, the x-axis is GMT. The estimates are based on 5-minute high frequency data on futures prices from January 1996 to December 2012 for S&P500 and FTSE100, and from January 1999 to December 2012 for Eurostoxx50.

The daily dynamic pattern for the time-varying threshold is captured by linking threshold value α_t with lagged values of estimated continuous volatility per 5 minute log-price change $\sigma_{t-1}/\sum_{i=1}^{n-1}\mathbb{I}_{\{|f_{t-1+i\Delta}-f_{t-1+(i-1)\Delta}|\leq\alpha_{t-1}\}}$. Taking both time-of-day factor and lagged continuous volatility I obtain formula for the time-varying threshold:

$$\alpha_{t,i} = 3\frac{\sigma_{t-1}}{\left(\sum_{i=1}^{n-1}\mathbb{I}_{\{|f_{t-1+i\Delta}-f_{t-1+(i-1)\Delta}|\leq\alpha_{t-1}\}}\right)^{0.49}}TOD_i$$

Appendix D - GMM conditions to estimate GPD and intensity parameters in the P measure

The aim of the GMM estimation of the P measure is to find the following vector of parameters for each tail:

$$\theta^P = [\alpha_0^{\pm}\bar{v}_{\psi}^{\pm}(tr^{\pm}); \alpha_1^{\pm}\bar{v}_{\psi}^{\pm}(tr^{\pm}); \xi^{\pm}; \sigma^{\pm}]$$

The four moments conditions are as follows:

$$\frac{1}{N}\sum_{t=1}^{N}\sum_{j=1}^{n-1}\phi_i^{\pm}\left(\psi^{\pm}(\Delta_j^{n,t}p) - tr^{\pm}\right)1_{\{\psi^{\pm}(\Delta_j^{n,t}p)>tr^{\pm}\}} = 0 \quad i = 1, 2$$

$$\frac{1}{N}\sum_{t=1}^{N}\sum_{j=1}^{n-1}1_{\{\psi^{\pm}(\Delta_j^{n,t}p)>tr^{\pm}\}} - \alpha_0^{\pm}\bar{v}_{\psi}^{\pm}(tr^{\pm}) - \alpha_1^{\pm}\bar{v}_{\psi}^{\pm}(tr^{\pm})CV_t = 0$$

$$\frac{1}{N}\sum_{t=2}^{N}\left(\sum_{j=1}^{n-1} 1_{\{\psi^{\pm}(\Delta_j^{n,t}p)>tr^{\pm}\}} - \alpha_0^{\pm}\bar{v}_{\psi}^{\pm}(tr^{\pm}) - \alpha_1^{\pm}\bar{v}_{\psi}^{\pm}(tr^{\pm})CV_t\right)CV_{t-1} = 0$$

where:

$$\phi_1^{\pm}(u) = -\frac{1}{\sigma^{\pm}} + \frac{\xi^{\pm}u}{(\sigma^{\pm})^2}\left(1+\frac{1}{\xi^{\pm}}\right)\left(1+\frac{\xi^{\pm}u}{\sigma^{\pm}}\right)^{-1}$$

$$\phi_2^{\pm}(u) = \frac{1}{(\xi^{\pm})^2}ln\left(1+\frac{\xi^{\pm}u}{\sigma^{\pm}}\right) - \frac{u}{\sigma^{\pm}}\left(1+\frac{1}{\xi^{\pm}}\right)\left(1+\frac{\xi^{\pm}u}{\sigma^{\pm}}\right)^{-1}$$

Appendix E - A short guide on how to get VRP(k) from the GMM estimates

This is a very short and basic instruction on how to derive VRP(k) for any given threshold k based on estimates. All of the following results are based on the derivations presented in the appendix of the Bollerslev and Todorov (2011b) paper.

Let us have a look at the tail volatility measure first. The measure can be presented as a sum of two components:

$$\int_{x>k} x^2 v(x)dx = 2\bar{v}_{\psi}^{+}(tr^{+}) * K_1 + k^2\bar{v}_{\psi}^{+}(e^k - 1)$$

The first part of the sum is directly determined by my estimates. For the selected threshold of $tr^{+} = 0.075$ I have estimated the value directly:

$$\bar{v}_{\psi}^{+}(tr^{+}) = \alpha_Q^{+}\bar{v}_{\psi}^{Q+}(0.075)$$

The multiplier K_1 is also directly defined by the estimated parameters:

$$K_1 = e^{-k/\xi^{+}}\xi^{+}\left(\frac{\xi^{+}}{\sigma^{+}}\right)^{-1/\xi^{+}}[\xi^{+}{}_3F_2\left(\frac{1}{\xi^{+}},\frac{1}{\xi^{+}},\frac{1}{\xi^{+}};1+\frac{1}{\xi^{+}},1+\frac{1}{\xi^{+}};\frac{\frac{\xi^{+}}{\sigma^{+}}(tr^{+}+1)-1}{e^k\frac{\xi^{+}}{\sigma^{+}}}\right)$$
$$+k_2F_1\left(\frac{1}{\xi^{+}},\frac{1}{\xi^{+}};1+\frac{1}{\xi^{+}};\frac{\frac{\xi^{+}}{\sigma^{+}}(tr^{+}+1)-1}{e^k\frac{\xi^{+}}{\sigma^{+}}}\right)]$$

The second part of the sum can be obtained from the approximation to the GPD. Following Bollerslev and Todorov (2011b) I assume that for a large threshold value the following approximation holds with equality:

$$1-\frac{\bar{v}_{\psi}^{+}(u+x)}{\bar{v}_{\psi}^{+}(x)} = G(u;\sigma^{+},\xi^{+})$$

where $G()$ denotes a GPD. Assuming that $x = tr^{+}$, $u = e^k - 1 - tr^{+}$ and $tr^{+} = 0.075$, it is quite straight forward that:

$$\bar{v}_{\psi}^{+}(e^k-1) = \left[1-G\left(e^k-1-tr+;\sigma^{+},\xi^{+}\right)\right]\bar{v}_{\psi}^{+}\left(tr^{+}\right)$$

Appendix F - Q-measure estimates of the right tail

Table 13: Q-measure: estimation results for the right tail

	S&P500	Eurostoxx50	FTSE100
ξ	0.1530	0.1143	0.1015
	(0.0115)	(0.0114)	(0.0122)
σ	0.0278	0.0329	0.0272
	(0.0006)	(0.0006)	(0.0004)
αv	0.8049	1.1443	0.7383
	(0.0184)	(0.0258)	(0.0156)

Notes: Table reports estimated parameters of the generalized Pareto distribution of the right-tail under the risk neutral Q-measure: ξ is the estimate of the shape parameter and σ is the estimate of the scale parameter. αv is the estimate of the average annualized jump intensity of +7.5% jump in the price level. The estimates are based on S&P500 and FTSE100 options data from January 1996 to December 2013 and Eurostoxx50 options data from March 2002 to December 2013. The log-moneyness of options used to estimate parameters were 1.1000, 1.0875 and 1.0750. Estimated standard errors are reported in parentheses.

Table 14: Q-measure: annualized jump intensity estimates for the right tail.

Jump Size	S&P500	Eurostoxx50	FTSE100
>7.5%	0.8049	1.1443	0.7383
>10%	0.3462	0.5523	0.3065
>20%	0.0262	0.0488	0.0170

Notes: Table reports annualized average jump intensities under the Q measure i.e. implied by the option prices. Jump sizes are in terms of percentage changes in price levels. In the case of S&P500 and FTSE100, averages are calculated from January 1996 to December 2013, and for Eurostoxx50 averages are calculated from March 2002 to December 2013. All the reported figures are based on generalized Pareto distribution estimates reported in Table 13.

Appendix G - P-meassure estimates for the right tail

Table 15: P-meassure: estimation results for the right tail

	S&P500	Eurostoxx50	FTSE100
ξ	0.2088	0.1648	0.2218
	(0.0671)	(0.0739)	(0.0415)
100σ	0.1834	0.1955	0.1714
	(0.0161)	(0.0189)	(0.0092)
α_0	-0.0020	-0.0013	-0.0028
	(0.0001)	(0.0001)	(0.0001)
α_1	0.0396	0.0291	0.0402
	(0.0006)	(0.0005)	(0.0006)

Notes: Table reports estimated parameters of the generalized Pareto distribution of the right-tail under the physical P-measure: ξ is the estimate of the shape parameter and σ is the estimate of the scale parameter. α_0 and α_1 are estimates of parameters of equation 11 linking jump intensities to the time-varying continuous volatilities. The estimates are based on high-frequency 5-minute futures prices from January 1996 to December 2012 for S&P500 and FTSE100, and from January 1999 to December 2012 for Eurostoxx50. Estimated standard errors are reported in parentheses.

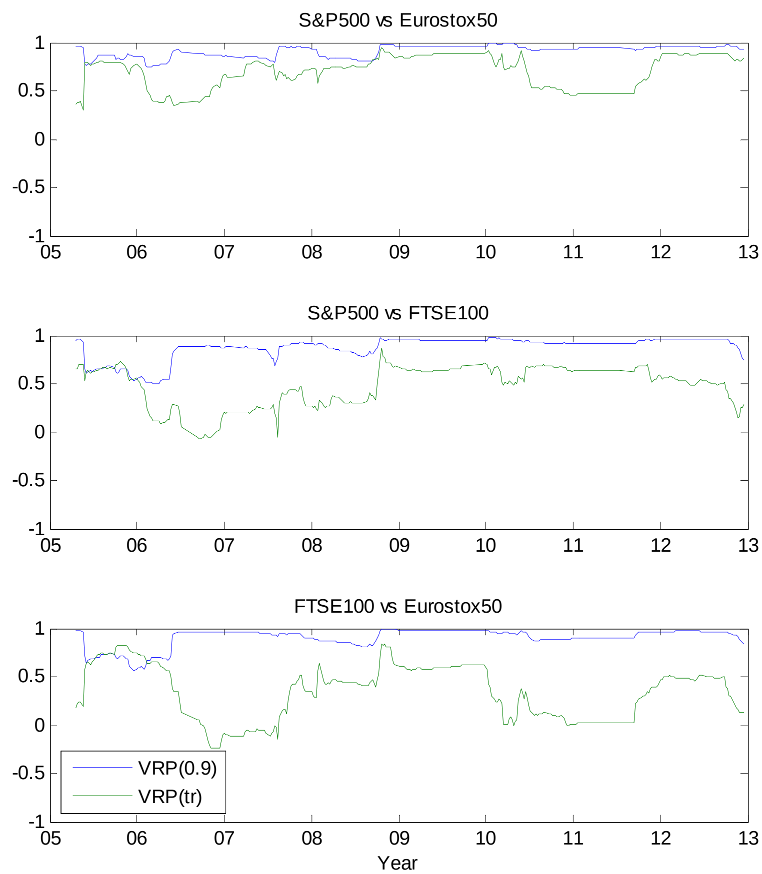
Table 16: P-measure: annualized jump intensity estimates for the right tail.

Jump Size	S&P500	Eurostoxx50	FTSE100
>7.5%	0.0062	0.0016	0.0187
>10%	0.0016	0.0003	0.0052
>20%	0.0001	0.0000	0.0002

Notes: Table reports annualized average jump intensities under the P measure i.e. based on the high frequency data estimation. Jump sizes are in terms of percentage changes in price levels. In case of S&P500 and FTSE100 averages are calculated from January 1996 to December 2012, and for Eurostoxx50 averages are calculated from January 1999 to December 2012. All the reported figures are based on generalized Pareto distribution estimates reported in Table 13.

Appendix H - Rolling Window correlations

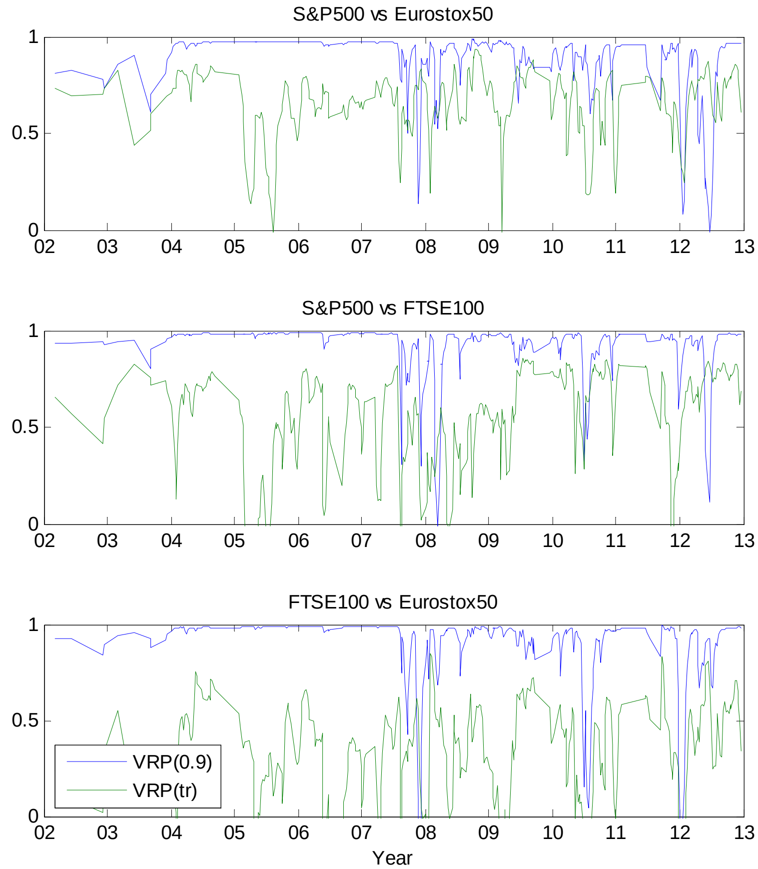
Figure 8: Time varying correlations of VRP(tr) and VRP(0.9) for different index pairs



Notes: Figures show time patterns of the 50-week rolling window r-Pearson correlation coefficients between different indices for crash-risk premia VRP(0.9) and non-crash-risk premia VRP(tr). Correlation coefficients are calculated on a common sample of weekly data for all three indices from March 2002 to December 2012.

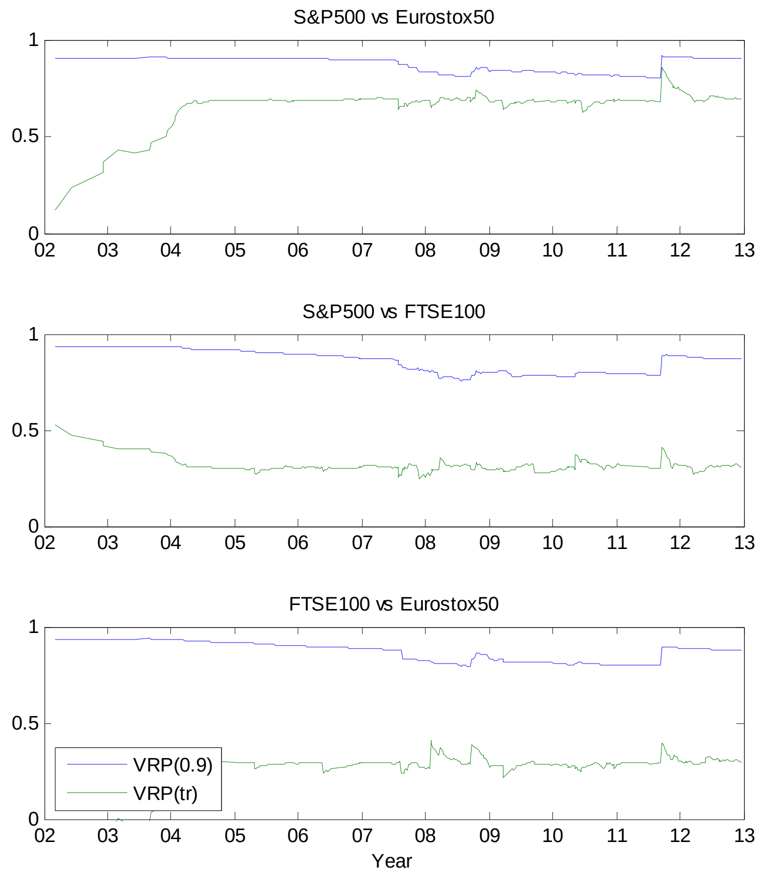
Appendix I - Dynamic correlations with different level equations

Figure 9: Dynamic conditional correlations on pure de-meaned data



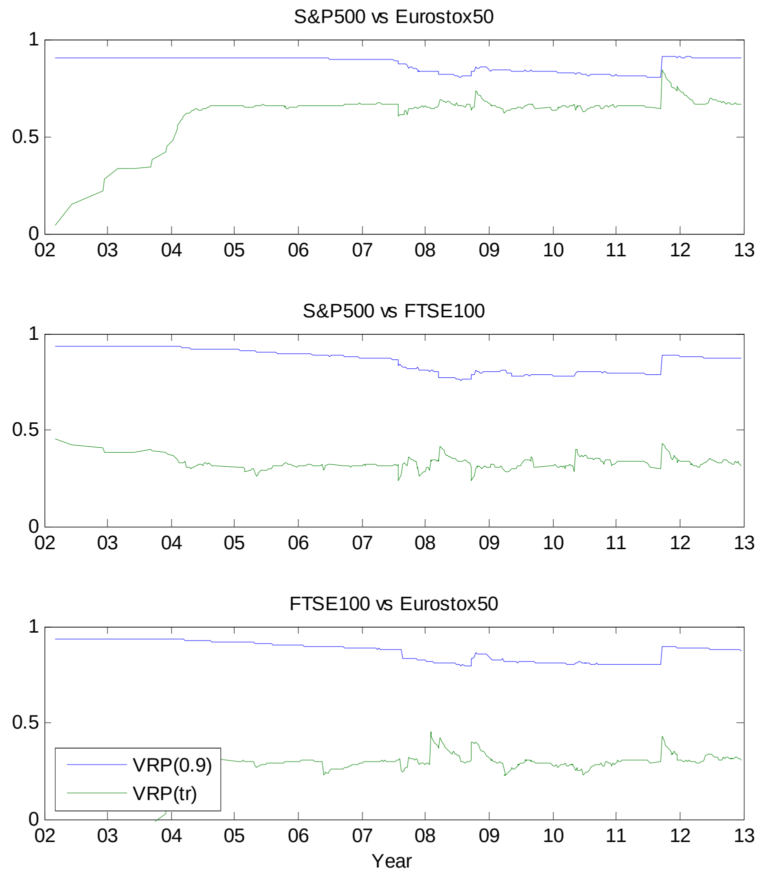
Notes: Figures show time patterns of the dynamic correlations of three index pairs for two premia measures: crash-risk VRP(0.9) (i.e. the premium for holding volatility risk associated with -10% jump in the price of the underlying index futures) and non-crash-risk VRP(tr) (i.e. the premium for holding volatility risk not related to the market crash). Dynamic correlations are calculated using Dynamic Conditional Correlation model of Engle (2002). The model is based on a common sample of weekly data for all three indices from March 2002 to December 2012. The level equation of either VRP(0.9) or VRP(tr) for all three indices is modelled as a constant, i.e. each level equation only de-means the data and does not account for any individual index persistence.

Figure 10: Dynamic conditional correlations, where the level equation is modelled as an AR(1) proces



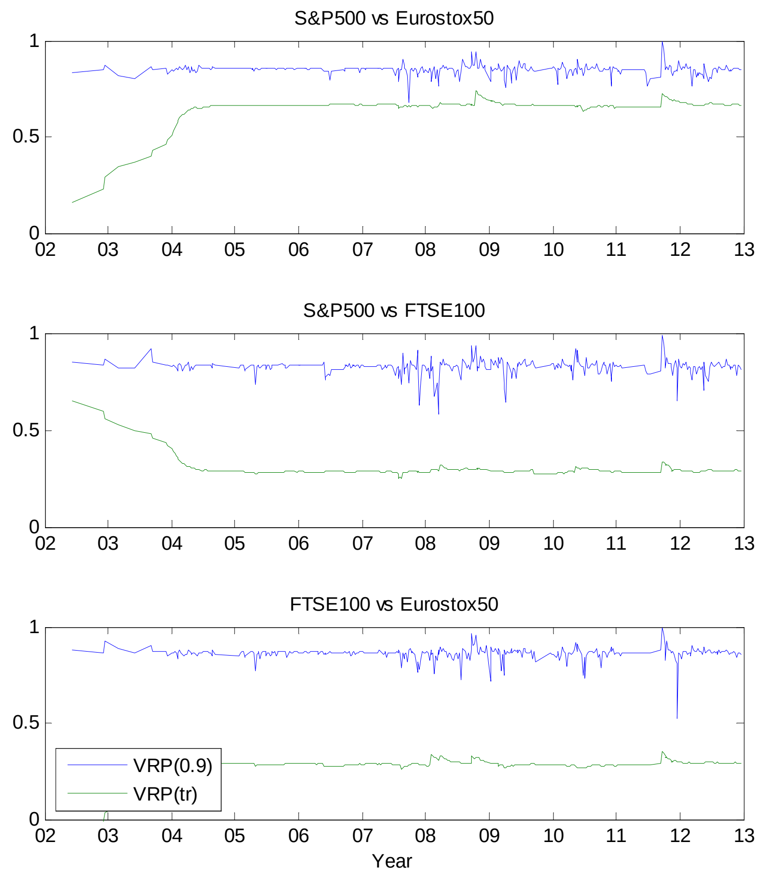
Notes: Figures show time patterns of the dynamic correlations of three index pairs for two premia measures: crash-risk VRP(0.9) (i.e. the premium for holding volatility risk associated with -10% jump in the price of the underlying index futures) and non-crash-risk VRP(tr) (i.e. the premium for holding volatility risk not related to the market crash). Dynamic correlations are calculated using Dynamic Conditional Correlation model of Engle (2002). The model is based on a common sample of weekly data for all three indices from March 2002 to December 2012. The level equation of either VRP(0.9) or VRP(tr) for all three indices is modelled individually as an AR(1) process.

Figure 11: Dynamic conditional correlations, where the level equation is modelled as an AR(SIC) proces



Notes: Figures show time patterns of the dynamic correlations of three index pairs for two premia measures: crash-risk VRP(0.9) (i.e. the premium for holding volatility risk associated with -10% jump in the price of the underlying index futures) and non-crash-risk VRP(tr) (i.e. the premium for holding volatility risk not related to the market crash). Dynamic correlations are calculated using Dynamic Conditional Correlation model of Engle (2002). The model is based on a common sample of weekly data for all three indices from March 2002 to December 2012. The level equation of either VRP(0.9) or VRP(tr) for all three indices is modelled individually as an AR(SIC) process, where the number of lags is selected using Bayesian information criterion.

Figure 12: Dynamic conditional correlations, where the level equation is modelled as a VAR(1) proces



Notes: Figures show time patterns of the dynamic correlations of three index pairs for two premia measures: crash-risk VRP(0.9) (i.e. the premium for holding volatility risk associated with -10% jump in the price of the underlying index futures) and non-crash-risk VRP(tr) (i.e. the premium for holding volatility risk not related to the market crash). Dynamic correlations are calculated using Dynamic Conditional Correlation model of Engle (2002). The model is based on a common sample of weekly data for all three indices from March 2002 to December 2012. The level equation of either VRP(0.9) or VRP(tr) for all three indices is modelled jointly as a VAR(1) process.