Sampling Properties of Contagion Tests*

Mardi Dungey[%+], Renée Fry[+%],
Brenda González-Hermosillo[*] and Vance L. Martin[#]

[+] CERF, University of Cambridge
[%]CAMA, Australian National University
[*]International Monetary Fund
[#]University of Melbourne

October 4, 2005

Abstract

The finite sample properties of various tests of contagion are investigated using a range of Monte Carlo experiments. The tests considered are the Forbes and Rigobon (FR) adjusted correlation test, the Favero and Giavazzi (FG) outlier test, the Pesaran and Pick (PP) threshold test, and the Bae, Karolyi and Stulz (BKS) co-exceedance test. Some variations of these tests are also investigated to correct for size distortions and weak instrument problems. The experiments focus on potential biases in testing for contagion, problems of spurious contagious linkages, the identification and measurement of common factors, the effects of alternative filtering methods, the importance of structural breaks, and the role of weak instruments. The finite sample results show that the FR and PP tests tend to be conservative tests as they are unlikely to find evidence of contagion when it does exist. In contrast, the FG and BKS tests are oversized as these tests tend to be more likely to find contagion when it does not exist. Two proposed variants of the Forbes and Rigobon test are shown to improve upon the sampling properties of this test. The PP and FG tests are shown to be affected by weak instrument and filtering problems. Finally, the change in the correlation coefficient test is shown to be a non-monotonic function of the strength of contagion for certain parameterisations resulting in non-monotonic power functions.

*We are grateful for comments from Monica Billio, Sam Ouliaris, Adrian Pagan, and from seminar participants at the University of Manchester, IMF and NZESG, Auckland 2004.

1 Introduction

There now exists a range of statistical procedures to test for financial market contagion.[1] Four prominent tests of contagion recently proposed are the Forbes and Rigobon (2002) adjusted correlation test (FR), the Favero and Giavazzi (2002) outlier test (FG), the Pesaran and Pick (2004) threshold test (PP), and the Bae, Karolyi and Stulz (2003) co-exceedance test (BKS).[2] Implementation of these tests involves a number of empirical and modelling issues, including the dating of crisis periods, choice of filtering methods to identify shocks, identification of common factors, simultaneity bias and weak instruments, and the ability of test statistics to condition on structural breaks. Table 1 provides a checklist of these properties. All of these issues have implications for the sampling properties of the test statistics in terms of their size and power to identify contagion. A manifestation of the differences in the statistical properties of these tests is in their empirical implementation whereby the conclusions can differ markedly.

The aim of this paper is to investigate the finite sampling properties of these four tests of contagion by performing an extensive number of Monte Carlo experiments.[3] Both size and power comparisons are performed under various scenarios that include increases in asset return volatility arising from both contagion and structural breaks. Implications of including autocorrelation and GARCH conditional volatility structures in the data generating process (DGP) are also examined.

Apart from the four contagion tests already mentioned, some variations of these tests are also investigated in the Monte Carlo experiments. Two variants of the Forbes and Rigobon adjusted correlation test are proposed to correct the asymptotic variance and hence the size of the test, where overlapping data between crisis and noncrisis periods are used. A multivariate version of the Forbes and Rigobon test proposed by Dungey, Fry, González-Hermosillo and

[1] For a discussion of the definition of contagion, see Pericoli and Sbracia (2003), the World Bank website on contagion and Dornbusch, Park and Claessens (2000).

[2] Although this list of tests is by no means exhaustive, it does perhaps represent the main tests used in applied work with other tests providing important extensions or representing special cases. Some examples of related tests are the DCC test of Rigobon (2003) which is a multivariate extension of the Forbes and Rigobon test (2002), the probability model test of Eichengreen, Rose and Wyplosz (1995, 1996) which is a special case of the Bae, Karolyi and Stulz (2003) co-exceedance test, and more recently the test of Bekaert, Harvey and Ng (2005) which emphasises the role of conditional volatility in modelling contagion. Other approaches include Bayesian methods, Dasgupta (2002), Ciccarelli and Rebucci (2004), and Markov switching approaches such as Jeanne and Masson (2000).

[3] There has been little work investigating the properties of alternative tests of contagion. Some exceptions are Forbes and Rigobon (2002), Pesaran and Pick (2004), and Walti (2003).

Martin (2005b) is also investigated to detect spurious contagious transmission mechanisms. To identify potential weak instrument problems with the Pesaran and Pick test, the results based on both IV and OLS estimators are examined. Finally, the refinements to the Bae, Karolyi and Stulz (2003) test suggested in Dungey, Fry, González-Hermosillo and Martin (2005a), are also adopted.

Some of the key results of the experiments are that the Forbes and Rigobon test and the Pesaran and Pick test are undersized. In contrast the Favero and Giavazzi test and the Bae, Karolyi and Stulz test are found to be oversized. The proposed adjustments to the Forbes and Rigobon test are shown to correct the size bias in the original version of this test. For certain parameterisations, the power functions of tests based on differences in correlation coefficients are shown to be nonmonotonic. The simulation results also show that incorrect diagnosis of structural breaks in testing for contagion can badly bias the test statistics towards the hypothesis of contagion. One solution to this problem is to use the multivariate version of the Forbes and Rigobon test which is found to have the best robustness properties in the presence of structural breaks.

The rest of the paper proceeds as follows. In Section 2 some background statistics describing the nature and magnitude of various financial crises are presented to motivate the structure of the data generating process (DGP) used in the Monte Carlo experiments. Section 3 provides some preliminary analysis of the ability of the DGP to simulate financial crises. An overview of the alternative testing procedures investigated in the paper are discussed in Section 4. Section 5 provides the finite sampling properties of these testing procedures, while concluding comments are contained in Section 6.

2 Background

To identify some of the key characteristics of financial crises that are used in parameterizing the DGP in the Monte Carlo experiments conducted below, Table 2 gives the covariances and correlations of daily returns for three financial crises. Each sample is divided into a noncrisis and crisis period, and statistics are computed for each of these as well as the total period. The three crises are the Tequila effect of December 1994 to March 1995, the Asian flu of July 1997 to March 1988 and the Russian cold of August 1998 to November 1998. The data presented not only cover three different regions, but also cover three different financial markets; namely, equity, currency and bond markets.

Table 1:

Summary of key issues in testing for contagion.

	Characteristic	FR	FG	PP	BKS
1	**Dating Crises**				
	A priori dating	✓			
	Endogenous dating		✓	✓	✓
2	**Filtering methods**				
	Slope dummy variables	✓			
	Intercept dummy variables		✓	✓	✓
3	**Controlling for common factors**				
	Observable variables	✓			✓
	Lagged variables	✓	✓	✓	
4	**Weak instruments**				
	Simultaneity bias	✓			✓
	Instrumental variable bias		✓	✓	
5	**Structural breaks**				
	Idiosyncratic break	✓	✓		
	Common break	✓	✓		✓
	Time-varying volatility	✓			

Inspection of the noncrisis and crisis covariance matrices shows that all crises are characterized by very large increases in volatility. The extent of the increase in volatility during the crisis periods is further highlighted in Figure 1 which gives time series plots of the financial returns of various countries in each of the three financial crises. Associated with increases in both variances and covariances during the crisis periods, are increases in correlations in most cases. A counter example is the correlation between Bulgaria and Russia during the Russian crisis where the correlation actually falls during the crisis period. Extending the comparison to between crisis and total correlations shows that the correlation between Indonesia and Korea during the Asian crisis fell marginally from 0.027 (total) to 0.025 (crisis).[4]

The increase in volatility during the crisis periods highlighted in Table 2 can be attributed to an increase in the volatility of either the economic and financial factors that jointly underlie financial returns (common factors), or factors that are specific to particular financial market (idiosyncratic factors), or the result of an additional factor over and above the common and idiosyncratic factors that are specific to crises per se (contagion), or a combination of all channels. A key issue confronting tests of contagion concerns identifying these three broad channels of transmission.

A further issue that is important in testing for contagion relates to the duration of crisis periods. Crisis periods tend to be relatively short. For example the Tequila and Russian examples in Table 2 are 54 and 62 days respectively. The short nature of these crises suggests that asymptotic distribution theory may not provide an accurate approximation to the finite sample distributions of statistics used to test for contagion.

3 Simulating Crises

In this section, a model of financial crises is developed by first outlining financial market linkages in a tranquil, noncrisis period, and then extending the model to include crisis period linkages. The crisis model is based on the framework of Dungey, Fry, González-Hermosillo and Martin (2002, 2005a), which, in turn, is motivated by the class of factor models commonly adopted in finance where the determinants of asset returns are decomposed into common factors and idiosyncratic factors (see also Pericoli and Sbracia (2003)). The model is couched in terms of the financial returns on assets in three financial markets, although the

[4] Other examples where noncrisis correlations are larger than crisis correlations are documented by Forbes and Rigobon (2002) and Baig and Goldfajn (1999).

Table 2:

Variances (diagonal), covariances (lower triangle) and correlations (upper triangle) of financial returns for selected crises: T_x and T_y are the number of observations in the noncrisis and crisis periods respectively, and $T = T_x + T_y$ represents the total number of observations in the sample.

Country	Noncrisis			Crisis			Total		
				Tequila effect (equity markets)[(a)]					
	1/6/94 to 11/12/94 ($T_x = 143$)			12/12/94 to 2/3/95 ($T_y = 54$)			1/6/94 to 2/3/95 ($T = 197$)		
Arg.	1.982	0.387	0.288	16.752	0.755	0.449	6.194	0.607	0.423
Chi.	0.411	0.567	0.116	3.765	1.484	0.279	1.374	0.828	0.226
Mex.	0.553	0.119	1.853	6.340	1.172	11.890	2.274	0.444	4.669
				Asian Flu (currency markets)[(b)]					
	3/3/97 to 3/7/97 ($T_x = 84$)			4/7/97 to 31/3/98 ($T_y = 190$)			3/3/97 to 31/3/98 ($T = 274$)		
Ind.	0.016	−0.197	0.136	35.070	0.125	0.521	24.373	0.127	0.445
Kor.	−0.004	0.024	−0.001	2.718	13.446	0.110	1.908	9.325	0.094
Tha.	0.036	0.000	4.314	6.948	0.910	5.079	4.823	0.630	4.828
				Russian Cold (bond markets)[(c)]					
	12/2/98 to 16/8/98 ($T_x = 119$)			17/8/98 to 15/11/98 ($T_y = 62$)			12/2/98 to 15/11/98 ($T = 181$)		
Bul.	0.101	0.367	0.713	3.100	0.586	0.549	1.117	0.558	0.556
Pol.	0.013	0.013	0.292	0.385	0.139	0.559	0.139	0.055	0.522
Rus.	0.216	0.031	0.904	3.624	0.783	14.064	1.366	0.286	5.404

(a) Equity returns are computed as the difference of the natural logarithms of daily share prices, expressed as a percentage.

(b) Currency returns are computed as the difference of the natural logarithms of daily bilateral exchange rates, defined relative to the US dollar, expressed as a percentage.

(c) Daily change in bond yields, expressed as a percentage.

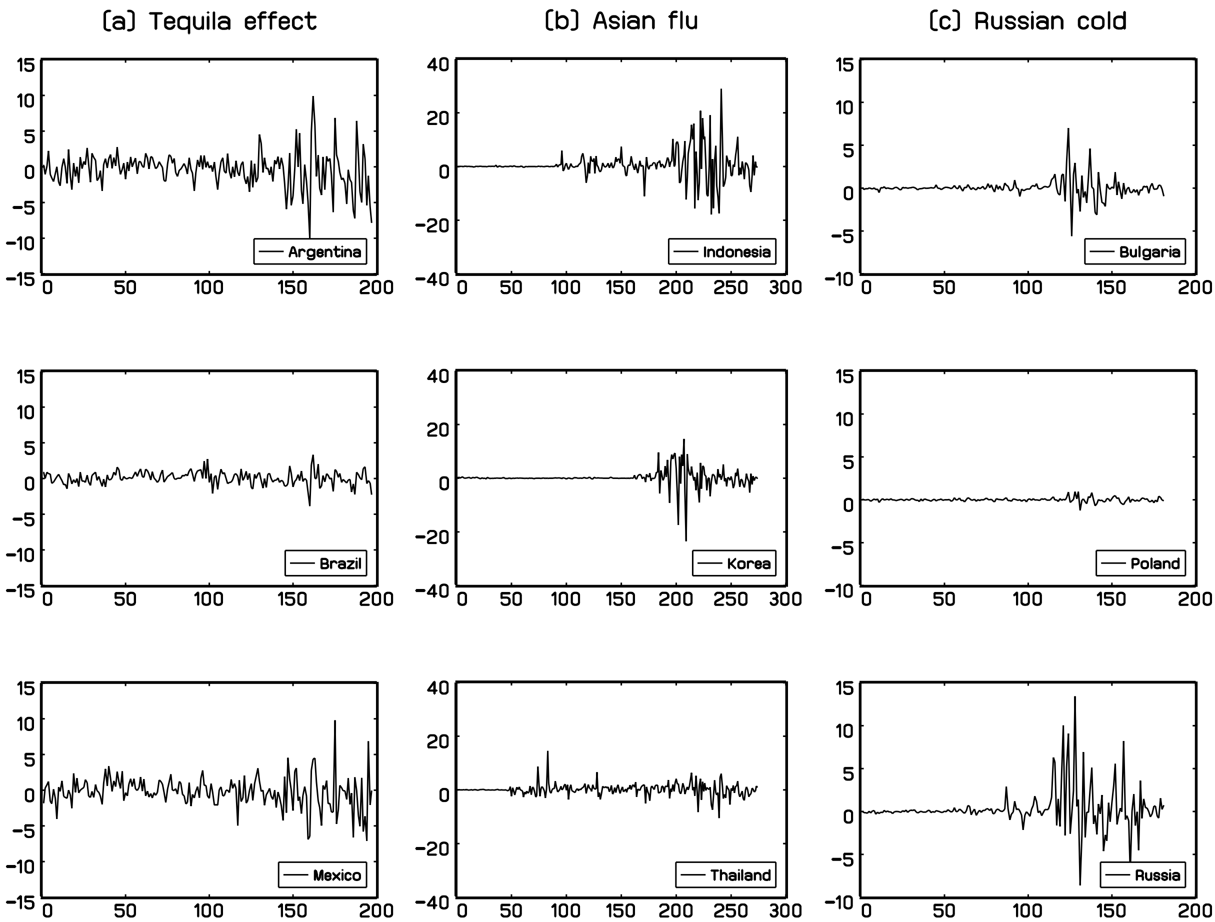

Figure 1: Empirical examples of selected financial crises: Tequila effect based on equity returns (%) (June 1994 - March 1995), Asian flu based on currency returns (%) (March 1997 - March 1998), Russian cold based on changes in bond yields (%) (August 1998 - November 1998).

analysis can easily be extended to include more markets. Increases in volatility arise from two transmission mechanisms in the model: structural shifts in the common and idiosyncratic factors, and contagion. This model serves as the basis for the DGP used in the Monte Carlo experiments conducted in Section 4. An important feature of the model is that it highlights special features in the data which are important in understanding the properties of contagion tests.

3.1 The Noncrisis Model

The noncrisis model consists of a one (common) factor model where returns $(x_{i,t})$ are a function of a common factor (w_t) and an idiosyncratic component $(u_{i,t})$

$$x_{i,t} = \lambda_i w_t + \phi_i u_{i,t} \qquad i = 1, 2, 3, \tag{1}$$

where

$$w_t \sim N(0,1) \tag{2}$$

$$u_{i,t} \sim N(0,1) \qquad i = 1, 2, 3, \tag{3}$$

are assumed to be independent initially. The common factor captures systemic risk which impacts upon asset returns with a loading of λ_i. The idiosyncratic components capture unique aspects to each return, and impact upon asset returns with a loading of ϕ_i. In a noncrisis period, the idiosyncratics represent potentially diversifiable non-systemic risk. In the special case where $\lambda_1 = \lambda_2 = \lambda_3 = 0$, the markets are segmented with volatility in asset returns entirely driven by their respective idiosyncratic factors. The assumption that the common and idiosyncratics are identically distributed can be relaxed by including autocorrelation and conditional volatility in the form of GARCH; see for example Dungey, Martin and Pagan (2000), Dungey and Martin (2004) and Bekaert, Harvey and Ng (2005).

3.2 The Crisis Model

The crisis model is an extension of the noncrisis model in (1) to (3) by allowing for a structural break in the world factor, as well as for increases in asset return volatility resulting from an additional propagation mechanism caused by contagion. This model is further extended in Section 4 to allow for structural breaks in the idiosyncratic factors. To distinguish the crisis period from the noncrisis period, returns in the crisis period are denoted as $y_{i,t}$. Contagion is assumed to transmit from country 1 to the remaining two countries, countries 2 and 3. Additional dynamics can be included by allowing for contagious feedback effects, provided that there are sufficient identifying restrictions to be able to determine each propagation mechanism.

The factor structure during the crisis period is specified as

$$\begin{aligned} y_{1,t} &= \lambda_1 w_t + \phi_1 u_{1,t} && (4)\\ y_{2,t} &= \lambda_2 w_t + \phi_2 u_{2,t} + \delta_2 \phi_1 u_{1,t} && (5)\\ y_{3,t} &= \lambda_3 w_t + \phi_3 u_{3,t} + \delta_3 \phi_1 u_{1,t}, && (6) \end{aligned}$$

where

$$\begin{aligned} w_t &\sim N\left(0, \omega^2\right) && (7)\\ u_{i,t} &\sim N\left(0, 1\right), \qquad i = 1, 2, 3. && (8) \end{aligned}$$

Contagion is defined as shocks originating in country 1, $\phi_1 u_{1,t} = y_{1,t} - \lambda_1 w_t$, which impact upon the asset returns of countries 2 and 3, over and above the contribution of the systemic factor $(\lambda_i w_t)$ and the country's idiosyncratic factor $(\phi_i u_{i,t})$. The strength of the contagion channel is determined by the parameters δ_2 and δ_3, in (5) and (6) for countries 2 and 3 respectively. The transmission of volatility through the systemic factor is captured by the structural break in the common factor (w_t), given by equation (7), where the variance in the common factor increases from unity in the non-crisis period to $\omega^2 > 1$, in the crisis period. For example, it may capture the impact of changes in trading volumes or changes in the general risk profile of investors. An implication of the model is that contagion adds to the level of nondiversifiable risk when diversification is needed most. This is highlighted by equations (4) to (6) as the idiosyncratic of $y_{1,t}$, given by $u_{1,t}$, now represents a common factor (nondiversifiable) during the crisis period; see also Walti (2003) for further discussion of this point.

Let asset returns over the total period be denoted as $z_{i,t}$, which are given by concatenating the noncrisis and crisis period returns. Letting the sample periods of the noncrisis and crisis periods be T_x and T_y respectively, then

$$z_{i,t} = \left(x_{i,1}, x_{i,2}, \cdots, x_{i,T_x}, y_{i,T_x+1}, y_{i,T_x+2}, \cdots, y_{i,T_x+T_y}\right)', \qquad (9)$$

represents the full sample of asset returns for the i^{th} country. The dynamics of the common factor over the total period are summarized as

$$w_t \sim \left\{ \begin{array}{lll} N(0,1) & : & t = 1,2,\cdots,T_x \\ N(0,\omega^2) & : & t = T_x+1, T_x+2, \cdots, T_x+T_y \end{array} \right. .$$

3.3 Covariance Structure

The specified model in (1) to (8) captures some of the key empirical features of financial crises. To highlight these properties, consider the variance-covariance matrices of returns for the two sample periods. Using the independence assumption of the factors w_t and $u_{i,t}$, $i = 1,2,3$ in (2) and (3), the variance-covariance matrix during the noncrisis period is obtained from (1)

$$\Omega_x = \left[\begin{array}{ccc} \lambda_1^2+\phi_1^2 & \lambda_1\lambda_2 & \lambda_1\lambda_3 \\ \lambda_1\lambda_2 & \lambda_2^2+\phi_2^2 & \lambda_2\lambda_3 \\ \lambda_1\lambda_3 & \lambda_2\lambda_3 & \lambda_3^2+\phi_3^2 \end{array} \right]. \tag{10}$$

Similarly, using (4) to (8) the variance-covariance matrix during the crisis period is

$$\Omega_y = \left[\begin{array}{ccc} \lambda_1^2\omega^2+\phi_1^2 & \lambda_1\lambda_2\omega^2+\delta_2\phi_1^2 & \lambda_1\lambda_3\omega^2+\delta_3\phi_1^2 \\ \lambda_1\lambda_2\omega^2+\delta_2\phi_1^2 & \lambda_2^2\omega^2+\phi_2^2+\delta_2^2\phi_1^2 & \lambda_2\lambda_3\omega^2+\delta_2\delta_3\phi_1^2 \\ \lambda_1\lambda_3\omega^2+\delta_3\phi_1^2 & \lambda_2\lambda_3\omega^2+\delta_2\delta_3\phi_1^2 & \lambda_3^2\omega^2+\phi_3^2+\delta_2^2\phi_1^2 \end{array} \right]. \tag{11}$$

The proportionate increase in volatility of the source crisis country is obtained directly from (10) and (11)

$$\theta_1 = \frac{Var(y_{1,t})}{Var(x_{1,t})} - 1 = \frac{\lambda_1^2\omega^2+\phi_1^2}{\lambda_1^2+\phi_1^2} - 1 = \frac{\lambda_1^2(\omega^2-1)}{\lambda_1^2+\phi_1^2}. \tag{12}$$

When there is no structural break $(\omega = 1)$, there is no increase in volatility in the source country $(\theta_1 = 0)$. In this case, any increase in the volatility of asset returns in countries 2 and 3 is solely the result of contagion, $\delta_i > 0, i = 2,3$. The general expressions for the proportionate increases in volatility in countries 2 and 3 are respectively given by

$$\theta_2 = \frac{Var(y_{2,t})}{Var(x_{2,t})} - 1 = \frac{\lambda_2^2\omega^2+\phi_2^2+\delta_2^2\phi_1^2}{\lambda_2^2+\phi_2^2} - 1 = \frac{\lambda_2^2(\omega^2-1)+\delta_2^2\phi_1^2}{\lambda_2^2+\phi_2^2} \tag{13}$$

$$\theta_3 = \frac{Var(y_{3,t})}{Var(x_{3,t})} - 1 = \frac{\lambda_3^2\omega^2+\phi_3^2+\delta_3^2\phi_1^2}{\lambda_3^2+\phi_3^2} - 1 = \frac{\lambda_3^2(\omega^2-1)+\delta_3^2\phi_1^2}{\lambda_3^2+\phi_3^2}. \tag{14}$$

These expressions show that volatility can increase for two reasons: increased volatility in the systemic factor $\left(\lambda_i^2(\omega^2-1)\right)$ and increased volatility arising from contagion $\left(\delta_i^2\phi_1^2\right)$. A fundamental requirement of any test of contagion is to be able to identify the relative magnitudes of these two sources of increased volatility.

To highlight some of the features of the model further, Table 3 presents the variance-covariance matrices for the noncrisis (Ω_x) and crisis (Ω_y) periods for two experiments with the adopted parameterization given in the caption of the Table. The first experiment is where there is contagion ($\delta_i = 5$) and no structural break in the common factor ($\omega = 1$). Figure 2(a) contains simulated time series of asset returns for the three countries under this scenario. The sample sizes are $T_x = 100$ for the non-crisis period and $T_y = 50$ for the crisis period. Asset return variances in countries 2 and 3 increase by a factor of 869% and 500% respectively. These numbers are comparable to the increases in the variances of financial returns for the three historical financial crises reported in Table 2. By definition the volatility of country 1 does not change with its variance equal to 20 in both periods.

The case where there is additional volatility in market fundamentals during a crisis period ($\omega = 5$), is given by Experiment 2 in Table 3. Figure 2(b) contains simulated time series of asset returns for the three countries under this scenario. The sample sizes are as before, namely, $T_x = 100$ for the non-crisis period and $T_y = 50$ for the crisis period. Inspection of the variances of all three countries show enormous increases in volatility during the crisis period. Associated with the increases in the variances highlighted in Table 3, are also increases in the covariances. From (11), these increases in the covariances arise from the increase in volatility of the market fundamentals ($\lambda_i \lambda_j \omega^2$) and of course contagion $\left(\delta_i \delta_j \phi_1^2\right)$, $\forall i, j$ with $i \neq j$, and $\delta_1 = 1$.

4 Review of Contagion Tests

This section presents the details of eight tests of contagion whose size and power properties are investigated in the Monte Carlo experiments below. Four of the eight tests are commonly used in empirical work to test for contagion. The remaining tests investigated represent extensions of these four tests which are designed to correct for either potential biases arising from size distortions and weak instrument problems, or to provide extensions in order to expand the types of contagious linkages that can be tested. The set of tests are summarized in Table 4.

4.1 Forbes and Rigobon Adjusted Correlation Test (FR1)

Forbes and Rigobon (2002) identify contagion as an increase in the correlation of returns between noncrisis and crisis periods having adjusted for market fundamentals and any in-

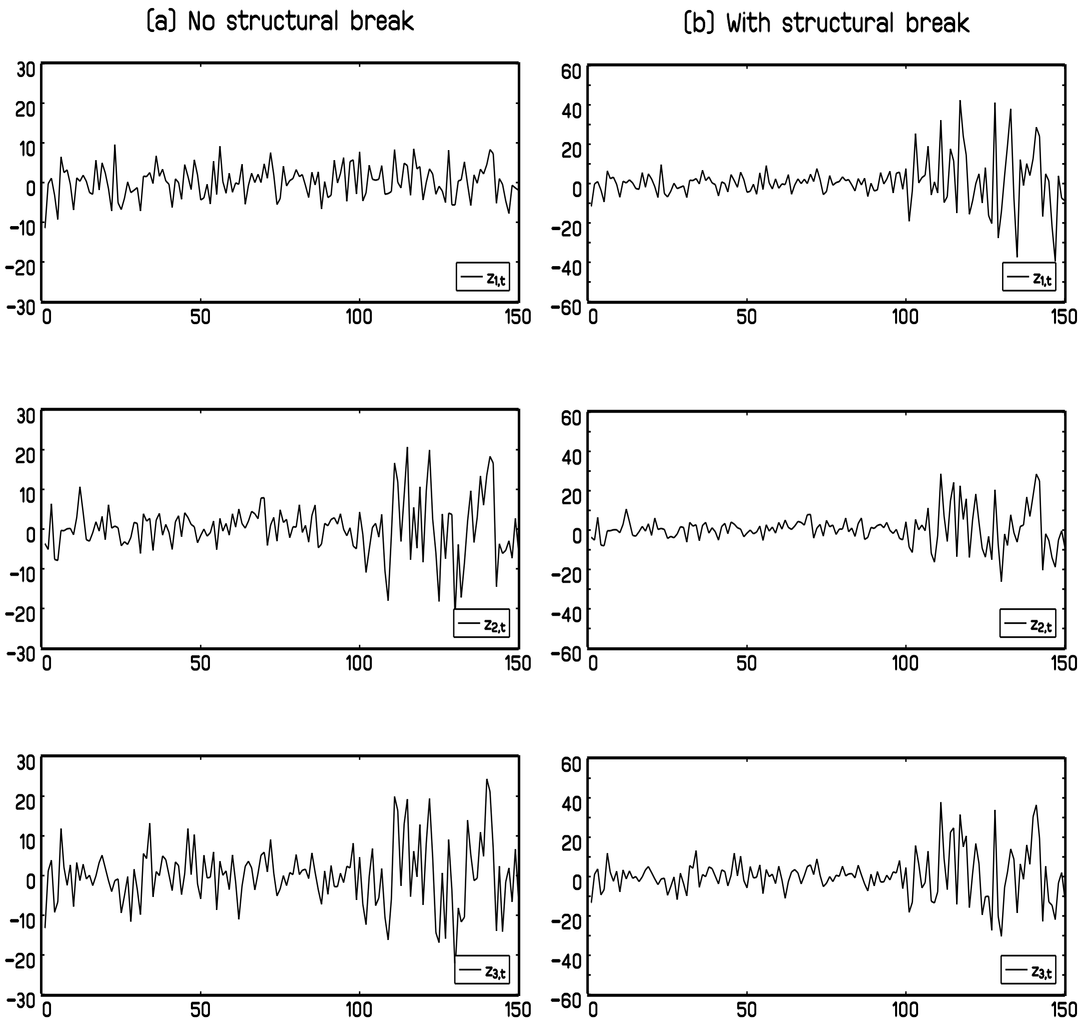

Figure 2: Simulated crisis data: (a) Contagion with no structural break in common factor; (b) Contagion with structural break in common factor. The crisis period is represented by the last 50 observations.

Table 3:

Noncrisis (Ω_x) and crisis (Ω_y) variance-covariance matrices for alternative parameterization. The DGP is based on (1) to (8), with common factor parameters $\lambda_1 = 4, \lambda_2 = 2, \lambda_3 = 3$; idiosyncratic parameters $\phi_1 = 2, \phi_2 = 10, \phi_3 = 4$; contagion parameters $\delta_2 = 5, \delta_3 = 5$; and common structural break parameters $\omega = 1, 5$. The sample sizes are $T_x = 100$ and $T_y = 50$.

Experiment 1: Without structural break ($\omega = 1$)

$$\Omega_x = \begin{bmatrix} 20 & 8 & 12 \\ 8 & 104 & 6 \\ 12 & 6 & 25 \end{bmatrix} \qquad \Omega_y = \begin{bmatrix} 20 & 28 & 32 \\ 28 & 204 & 106 \\ 32 & 106 & 125 \end{bmatrix}$$

Experiment 2: With structural break ($\omega = 5$)

$$\Omega_x = \begin{bmatrix} 20 & 8 & 12 \\ 8 & 104 & 6 \\ 12 & 6 & 25 \end{bmatrix} \qquad \Omega_y = \begin{bmatrix} 404 & 220 & 350 \\ 220 & 300 & 250 \\ 320 & 250 & 341 \end{bmatrix}$$

creases in volatility of the source country. The noncrisis period is typically taken as the pooled sample of returns $z_{i,t}$, in (9).

Let

$$\begin{aligned} \rho_z &= Corr\left(z_{i,t}, z_{j,t}\right) \\ \rho_y &= Corr\left(y_{i,t}, y_{j,t}\right), \end{aligned} \tag{15}$$

represent the correlations between the returns in country i and country j in the total (noncrisis) and crisis periods respectively. To test for contagion from country i to country j, the statistic is

$$FR1 = \frac{\frac{1}{2}\ln\left(\frac{1+\widehat{\nu}_y}{1-\widehat{\nu}_y}\right) - \frac{1}{2}\ln\left(\frac{1+\widehat{\rho}_z}{1-\widehat{\rho}_z}\right)}{\sqrt{\frac{1}{T_y-3} + \frac{1}{T_z-3}}}, \tag{16}$$

where

$$\widehat{\nu}_y = \frac{\widehat{\rho}_y}{\sqrt{1 + \left(\frac{s_{y,i}^2 - s_{z,i}^2}{s_{z,i}^2}\right)\left(1 - \widehat{\rho}_y^2\right)}}, \tag{17}$$

represents an adjusted correlation coefficient that takes into account increases in volatility in the source country (country i), $\widehat{\rho}_z$ and $\widehat{\rho}_y$ are estimators of the correlation coefficients in (15), and $s_{z,i}^2$ and $s_{y,i}^2$ are the sample variances corresponding to $\sigma_{z,i}^2$ and $\sigma_{y,i}^2$ respectively. The

Table 4:

Summary of alternative contagion tests.

Test		Description
FR1	:	Forbes and Rigobon test with the noncrisis period based on total sample period returns
FR2	:	Forbes and Rigobon test with the noncrisis period based on noncrisis sample period returns
FR3	:	Forbes and Rigobon test with the noncrisis period based on total sample period returns, but with an adjusted variance
FRM	:	Forbes and Rigobon multivariate test
FG	:	Favero and Giavazzi outlier test
PP1	:	Pesaran and Pick threshold test
PP2	:	Pesaran and Pick threshold test without simultaneity correction
BKS	:	Bae, Karolyi and Stulz co-exceedance test

statistic makes use of the Fisher transformation to improve the asymptotic approximation (Kendall and Stuart, Vol.1, p.390-391). Under the null hypothesis of no contagion from country i to country j, $\nu_y = \rho_z$, and

$$FR1 \xrightarrow{d} N(0,1). \tag{18}$$

4.2 Alternative Forbes and Rigobon Tests (FR2,FR3)

An alternative to the Forbes and Rigobon test $(FR1)$, is to model the noncrisis period using just the non-crisis returns, namely $x_{i,t}$. The test statistic is now obtained by replacing $z_{i,t}$ in (16) by $x_{i,t}$

$$FR2 = \frac{\frac{1}{2}\ln\left(\frac{1+\widehat{\nu}_y}{1-\widehat{\nu}_y}\right) - \frac{1}{2}\ln\left(\frac{1+\widehat{\rho}_x}{1-\widehat{\rho}_x}\right)}{\sqrt{\frac{1}{T_y-3} + \frac{1}{T_x-3}}}, \tag{19}$$

where

$$\widehat{\nu}_y = \frac{\widehat{\rho}_y}{\sqrt{1 + \left(\frac{s^2_{y,i} - s^2_{x,i}}{s^2_{x,i}}\right)\left(1 - \widehat{\rho}^2_y\right)}}, \tag{20}$$

$\widehat{\rho}_x$ is the estimator of the correlation coefficient $\rho_x = Corr(x_{i,t}, x_{j,t})$, and $s^2_{x,i}$ is the sample variance corresponding to $\sigma^2_{x,i}$. Under the null hypothesis of no contagion from country i to country j, $\nu_y = \rho_x$, and

$$FR2 \xrightarrow{d} N(0,1). \tag{21}$$

An advantage of using $FR2$ is that the assumption of independence between the crisis and noncrisis samples is satisfied. This is not the case with the $FR1$ statistic where the noncrisis period is defined as the total sample period. For this case the correct asymptotic variance is given by

$$\begin{aligned} Var(\widehat{\nu}_y - \widehat{\rho}_z) &= Var(\widehat{\nu}_y) + Var(\widehat{\rho}_z) - 2Cov(\widehat{\nu}_y, \widehat{\rho}_z) \\ &= Var(\widehat{\nu}_y) + Var(\widehat{\rho}_z) - 2Var(\widehat{\rho}_z) \\ &= Var(\widehat{\nu}_y) - Var(\widehat{\rho}_z) \\ &= \frac{1}{T_y} - \frac{1}{T_z}, \end{aligned} \tag{22}$$

where the second and third steps are based on the result $Cov(\widehat{\nu}_y, \widehat{\rho}_z) = Var(\widehat{\rho}_z)$,[5] whilst the last step uses the result that for *iid* normal samples the asymptotic approximation of

[5]Let u_t and v_t be two *iid* random variables. Define the first and second sub-samples as T_1 and T_2, and

the variance of the correlation coefficient is given by the inverse of the sample size (Kendall and Stuart (1973, p.307)). All details are given in the Appendix. Using (22) in (16) yields a third version of the Forbes and Rigobon test statistic

$$FR3 = \frac{\frac{1}{2}\ln\left(\frac{1+\widehat{\nu}_y}{1-\widehat{\nu}_y}\right) - \frac{1}{2}\ln\left(\frac{1+\widehat{\rho}_z}{1-\widehat{\rho}_z}\right)}{\sqrt{\frac{1}{T_y-3} - \frac{1}{T_z-3}}}. \tag{23}$$

As the variance of $FR3$ is smaller than $FR1$ in (16), the latter statistic is expected to yield smaller values on average resulting in this test being biased towards not rejecting the null of no contagion.

4.3 Forbes and Rigobon Multivariate Test (FRM)

Dungey, Fry, González-Hermosillo and Martin (2005b) show that an alternative to the Forbes-Rigobon test is to perform a Chow structural break test using dummy variables, where the dependent and independent variables are scaled by the respective noncrisis standard deviations. One advantage of this formulation is that it provides a natural extension of the bivariate approach to a multivariate framework that jointly models and tests all combinations of contagious linkages. A further advantage is that it is computationally more easy to implement than the multivariate extension based on the DCC test proposed by Rigobon (2003).

The Forbes and Rigobon multivariate contagion test investigated here is based on the

---

the total sample period as T. The covariance between x_t and y_t over the total sample period is

$$\begin{aligned} Cov\left(x_t, y_t\right) &= \frac{1}{T}\sum_{t=1}^{T}\left(x_t-\overline{x}\right)\left(y_t-\overline{y}\right) \\ &= \frac{1}{T}\left(\sum_{t=1}^{T}\left(x_t-\overline{x}\right)\left(y_t-\overline{y}\right)+\sum_{t=1}^{T}\left(x_t-\overline{x}\right)\left(y_t-\overline{y}\right)\right) \\ &= \frac{1}{T}\left(\sum_{t=1}^{T_1}\left(x_t-\overline{x}\right)\left(y_t-\overline{y}\right)+\sum_{t=T_1+1}^{T}\left(x_t-\overline{x}\right)\left(y_t-\overline{y}\right)\right) \\ &= \frac{T_1}{T}Cov\left(x_t, y_t\right)_1+\frac{T_2}{T}Cov\left(x_t, y_t\right)_2 \end{aligned}$$

where $Cov\left(x_t, y_t\right)_i$ is the covariance between x_t and y_t for sub-period $i = 1, 2$. Because of independence following set of regression equations

$$\begin{aligned}
\left(\frac{z_{1,t}}{\sigma_{x,1}}\right) &= \alpha_{1,0}+\alpha_{1,d}d_t+\alpha_{1,2}\left(\frac{z_{2,t}}{\sigma_{x,2}}\right)+\alpha_{1,3}\left(\frac{z_{3,t}}{\sigma_{x,3}}\right) \\
&\quad +\gamma_{1,2}\left(\frac{z_{2,t}}{\sigma_{x,2}}\right)d_t+\gamma_{1,3}\left(\frac{z_{3,t}}{\sigma_{x,3}}\right)d_t+\eta_{1,t} \\
\left(\frac{z_{2,t}}{\sigma_{x,2}}\right) &= \alpha_{2,0}+\alpha_{2,d}d_t+\alpha_{2,1}\left(\frac{z_{1,t}}{\sigma_{x,1}}\right)+\alpha_{2,3}\left(\frac{z_{3,t}}{\sigma_{x,3}}\right) \\
&\quad +\gamma_{2,1}\left(\frac{z_{1,t}}{\sigma_{x,1}}\right)d_t+\gamma_{2,3}\left(\frac{z_{3,t}}{\sigma_{x,3}}\right)d_t+\eta_{2,t} \\
\left(\frac{z_{3,t}}{\sigma_{x,3}}\right) &= \alpha_{3,0}+\alpha_{3,d}d_t+\alpha_{3,1}\left(\frac{z_{1,t}}{\sigma_{x,1}}\right)+\alpha_{3,2}\left(\frac{z_{2,t}}{\sigma_{x,2}}\right) \\
&\quad +\gamma_{3,1}\left(\frac{z_{1,t}}{\sigma_{x,1}}\right)d_t+\gamma_{3,2}\left(\frac{z_{2,t}}{\sigma_{x,2}}\right)d_t+\eta_{3,t},
\end{aligned} \tag{24}$$

where

$$d_t=\begin{cases}1: & t>T_x\\ 0: & \text{otherwise}\end{cases}, \tag{25}$$

represents a crisis dummy variable and $\eta_{i,t}$ are disturbance terms. Tests of contagion amount to testing the significance of the parameters $\gamma_{i,j}$. For example, to test for contagion from country 1 to country 2 the null hypothesis is $\gamma_{2,1}=0$. It is also possible to perform joint tests such as testing contagion from country 1 to both countries 2 and 3. The null hypothesis in this case is $\gamma_{2,1}=\gamma_{3,1}=0$.

4.4 Favero and Giavazzi Outlier Test (FG)

The Favero and Giavazzi (2002) test is similar to the multivariate version of the Forbes and Rigobon test in (24) as both procedures amount to testing for contagion using dummy variables. The FG approach consists of two stages: (a) Identification of outliers using a VAR; (b) Estimation of a structural model that incorporates the outliers of the previous step using dummy variables. The dummy variables are defined as

$$d_{i,t}=\begin{cases}1: & |v_{i,t}|>3\sigma_{v,i}\\ 0: & \text{otherwise}\end{cases} \tag{26}$$

where there is a unique dummy variable corresponding to each outlier, and $v_{i,t}$ are the residuals from a VAR that contains the asset returns of all variables in the system with respective variances $\sigma^2_{v,i}$. That is, a dummy variable is constructed each time an observation is judged extreme, $|v_{i,t}|>3\sigma_{v,i}$, with a one placed in the cell corresponding to the point in time when the extreme observation occurs, and zero otherwise. Let $d_{1,t}, d_{2,t}$ and $d_{3,t}$, represent the idiosyncratic sets of dummy variables for countries 1 to 3 respectively, and $d_{c,t}$, be the set of dummy variables that are classified as common to all asset markets ie the dummy variables that are equivalent. This last feature of the model is needed to circumvent multicollinearity problems from using two equivalent variables.

The next stage of the Favero and Giavazzi framework is to specify the following structural model containing the dummy variables over the full sample period

$$\begin{aligned} z_{1,t} &= \alpha_{1,0} + \alpha_{1,2} z_{2,t} + \alpha_{1,3} z_{3,t} + \theta_1 z_{1,t-1} \\ &\quad + \gamma_{1,1} d_{1,t} + \gamma_{1,2} d_{2,t} + \gamma_{1,3} d_{3,t} + \gamma_{1,c} d_{c,t} + \eta_{1,t} \\ \\ z_{2,t} &= \alpha_{2,0} + \alpha_{2,1} z_{1,t} + \alpha_{2,3} z_{3,t} + \theta_2 z_{2,t-1} \\ &\quad + \gamma_{2,1} d_{1,t} + \gamma_{2,2} d_{2,t} + \gamma_{2,3} d_{3,t} + \gamma_{2,c} d_{c,t} + \eta_{2,t} \\ \\ z_{3,t} &= \alpha_{3,0} + \alpha_{3,1} z_{1,t} + \alpha_{3,2} z_{2,t} + \theta_3 z_{3,t-1} \\ &\quad + \gamma_{3,1} d_{1,t} + \gamma_{3,2} d_{2,t} + \gamma_{3,3} d_{3,t} + \gamma_{3,c} d_{c,t} + \eta_{3,t}, \end{aligned} \tag{27}$$

where $\eta_{i,t}$ are structural disturbance terms. The $\gamma_{i,j}$ are vectors of parameters in general, with dimensions corresponding to the number of dummy variables in $d_{1,t}, d_{2,t}, d_{3,t}$ and $d_{c,t}$. Treating the dummy variables as predetermined, this model is just identified and can be conveniently estimated by FIML using an instrumental variables estimator. The instruments chosen are the three lagged returns, the constant and all dummy variables. Tests of contagion are based on testing the $\gamma_{i,j}$ $\forall i \neq j$ parameters.

4.5 Pesaran and Pick Threshold Test (PP1)

The Pesaran and Pick (2004) contagion test is similar to the approach of Favero and Giavazzi (2002) as both involve identifying outliers initially and then using the outliers in a structural model to test for contagion. One important difference is that Pesaran and Pick do not define a dummy variable for each outlier, but combine the outliers associated with each variable into a single dummy variable. Formally, the Pesaran and Pick test restricts the parameters on the Favero and Giavazzi dummy variables associated with the extremes of a particular asset return to be equal.

The specification of the dummy variable for the i^{th} asset return is given by

$$d_{i,t} = \begin{cases} 1: & |v_{i,t}| > \tau_i \\ 0: & \text{otherwise} \end{cases}, \tag{28}$$

where $v_{i,t}$ $i = 1, 2, 3$, are the residuals from a VAR that contains the asset returns of all variables in the system, and τ_i is a threshold which picks out the biggest 10% of outliers in asset return i.[6]

The structural model is specified as

$$\begin{aligned} z_{1,t} &= \beta_{1,0} + \theta_1 z_{1,t-1} + \gamma_{1,2} d_{2,t} + \gamma_{1,3} d_{3,t} + \eta_{1,t} \\ z_{2,t} &= \beta_{2,0} + \theta_2 z_{2,t-1} + \gamma_{2,1} d_{1,t} + \gamma_{2,3} d_{3,t} + \eta_{2,t} \\ z_{3,t} &= \beta_{3,0} + \theta_3 z_{3,t-1} + \gamma_{3,1} d_{1,t} + \gamma_{3,2} d_{2,t} + \eta_{3,t}, \end{aligned} \tag{29}$$

where $\eta_{i,t}$ are structural disturbance terms. Unlike Favero and Giavazzi, Pesaran and Pick treat the dummy variables as endogenous. A further difference in the two approaches is that Pesaran and Pick do not include own dummy variables, whereas Favero and Giavazzi do. The structural model in (29) is just identified and can be conveniently estimated by FIML using an instrumental variables estimator. The instruments chosen are the three lagged returns, and the constant. Tests of contagion are based on testing the parameters $\gamma_{i,j}$, $\forall i \neq j$.

4.6 Adjusted Pesaran and Pick Threshold Test (PP2)

An alternative version of the Pesaran and Pick contagion test investigated in the Monte Carlo experiments is to estimate (29) by OLS and not adjust for any simultaneity bias. This form of the test is motivated by the possibility of weak instruments and its consequences for testing for contagion. These issues are discussed further below.

4.7 Bae, Karolyi and Stulz Co-exceedance Test (BKS)

The co-exceedance test of Bae, Karolyi and Stulz (2003) is also based on the dummy variable expression in (28) to identify periods of contagion. An important difference between the BKS

[6] Other choices of the switch point of the dummy variable could be based on the approach of Favero and Giavazzi (2002) in (26), or the exchange market pressure index used by Eichengreen, Rose and Wyplosz (1995, 1996). Baur and Schulze (2002) consider an endogenous approach, while Bae, Karolyi and Stulz (2003) adopt an asymmetric approach and consider positive and negative extreme returns separately. The approach adopted here has the advantage that it circumvents potential problems in the simulations conducted below when no outliers are detected.

approach and the Pesaran and Pick approach however, is that the dependent variable is also transformed to a dummy variable.

In the BKS framework, the dummy variables are commonly referred to as exceedances. Two versions of the BKS framework are considered. The first is based on a trivariate framework where the aim is to perform a joint test of contagion from country 1 to countries 2 and 3. Define the polychotomous dummy variable between the returns of countries 2 and 3, as

$$e_{2,3,t} = \begin{cases} 0: & d_{2,t} = 0 \text{ and } d_{3,t} = 0 \\ 1: & d_{2,t} = 1 \text{ and } d_{3,t} = 0 \\ 2: & d_{2,t} = 0 \text{ and } d_{3,t} = 1 \\ 3: & d_{2,t} = 1 \text{ and } d_{3,t} = 1 \end{cases}. \tag{30}$$

Points in time when both countries 2 and 3 experience extreme returns, $e_{2,3,t} = 3$, are referred to as co-exceedances. Associated with each value of the polychotomous dummy variable $e_{2,3,t}$, is a probability

$$p_{j,t} = \Pr\left(e_{2,3,t} = j\right), \qquad j = 0, 1, 2, 3.$$

To test for contagion, the following multinomial logit model is estimated by maximum likelihood where the probabilities are parameterized by the logistic function

$$p_{j,t} = \frac{\exp\left(\mu_j + \gamma_j d_{1,t}\right)}{\Sigma_{k=0}^{3} \exp\left(\mu_j + \gamma_j d_{1,t}\right)}, \qquad j = 0, 1, 2, 3, \tag{31}$$

where the restriction $\mu_0 = \gamma_0 = 0$, is chosen as the normalization. The inclusion of the dummy variable $d_{1,t}$, from (28), represents the extreme returns of country 1, and forms the basis of the contagion tests from this country to the other two countries. The test of contagion from country 1 to country 2 is based on testing that $\gamma_1 = 0$. For non-zero values of γ_1, extreme returns in country 1 impact upon returns in country 2. To test for contagion from country 1 to country 3, the pertinent restriction to be tested is $\gamma_2 = 0$. A joint test of contagion from country 1 to countries 2 and 3, is given by testing the restriction $\gamma_3 = 0$, as a non-zero value of γ_3 corresponds to where extreme shocks in the returns of country 1 impact simultaneously on countries 2 and 3. To test for contagion in other directions, $d_{1,t}$ is replaced by a dummy variable representing extreme returns of another country, and $e_{2,3,t}$ in (30) is appropriately redefined.[7]

[7] Another way to perform bidirectional tests of contagion is to define the polychotomous dummy variable $e_{2,3,t}$ in (30), simply in terms of one variable; that is, define a binary dummy variable. This is also the approach adopted by Eichengreen, Rose and Wyplosz (1995, 1996), except that they specify the underlying distribution to be normal, resulting in the use of a probit model.

4.8 Discussion

4.8.1 Biasedness in Testing for Contagion Using Correlations

An important feature of the FR2 test is that it is based on testing the difference in bivariate correlations between noncrisis and crisis periods. Using the expressions for the variance-covariance matrices in (10) and (11), the difference in the two correlations between countries 1 and 2 is immediately given by

$$\rho_{y_{1,t},y_{2,t}} - \rho_{x_{1,t},x_{2,t}} = \frac{\lambda_1\lambda_2\omega^2 + \delta_2\phi_1^2}{\sqrt{\lambda_1^2\omega^2 + \phi_1^2}\sqrt{\lambda_2^2\omega^2 + \phi_2^2 + \delta_2^2\phi_1^2}} - \frac{\lambda_1\lambda_2}{\sqrt{\lambda_1^2 + \phi_1^2}\sqrt{\lambda_2^2 + \phi_2^2}}. \qquad (32)$$

Some insight into this expression is obtained by looking at what happens as the strength of contagion continuously increases (with δ_2 replaced by δ) and when there is no structural break in the common factor ($\omega = 1$)

$$\begin{aligned}\lim_{\delta\to\infty}\left(\rho_{y_{1,t},y_{2,t}} - \rho_{x_{1,t},x_{2,t}}\right) &= \lim_{\delta\to\infty}\left(\frac{\lambda_1\lambda_2 + \delta\phi_1^2}{\sqrt{\lambda_1^2 + \phi_1^2}\sqrt{\lambda_2^2 + \phi_2^2 + \delta^2\phi_1^2}}\right) - \frac{\lambda_1\lambda_2}{\sqrt{\lambda_1^2 + \phi_1^2}\sqrt{\lambda_2^2 + \phi_2^2}} \\ &= \lim_{\delta\to\infty}\left(\frac{\lambda_1\lambda_2/\delta + \phi_1^2}{\sqrt{\lambda_1^2 + \phi_1^2}\sqrt{\lambda_2^2/\delta^2 + \phi_2^2/\delta^2 + \phi_1^2}}\right) - \frac{\lambda_1\lambda_2}{\sqrt{\lambda_1^2 + \phi_1^2}\sqrt{\lambda_2^2 + \phi_2^2}} \\ &= \frac{1}{\sqrt{\lambda_1^2 + \phi_1^2}}\left(\phi_1 - \frac{\lambda_1\lambda_2}{\sqrt{\lambda_2^2 + \phi_2^2}}\right). \qquad (33)\end{aligned}$$

For relatively "high" levels of contagion, the correlation in the noncrisis period $\left(\rho_{x_{1,t},x_{2,t}}\right)$ can exceed the crisis period correlation $\left(\rho_{y_{1,t},y_{2,t}}\right)$ when

$$\phi_1 - \frac{\lambda_1\lambda_2}{\sqrt{\lambda_2^2 + \phi_2^2}} < 0,$$

or

$$1 + \left(\frac{\phi_2}{\lambda_2}\right)^2 < \left(\frac{\lambda_1}{\phi_1}\right)^2. \qquad (34)$$

Thus the key magnitudes for determining the relative magnitudes of the correlations between asset returns in the two sample periods, are the relative sizes of the loadings of the common factor (λ_i) to the idiosyncratic factor (ϕ_i) for the two asset returns.

Figure 3 gives the difference in the crisis correlation $\left(\rho_y\right)$ and noncrisis correlation (ρ_x) between selected pairs of countries based on (32) for alternative values of the contagion parameter $(\delta = \delta_2 = \delta_3)$ and the structural break parameter (ω), with the values of the remaining parameters given by

$$\lambda_1 = 4, \lambda_2 = 2, \lambda_3 = 3, \phi_1 = 2, \phi_2 = 10, \phi_3 = 4.$$

For the case where there is no structural break in the common factor $(\omega = 1)$, Figure 3(a) shows that the difference in the correlations between countries 1 and 2 increases monotonically initially as δ increases and eventually flattens out to a difference of 0.3. This result is to be expected as the expression in (34) is not satisfied for this parameterization: $1 + (\phi_2/\lambda_2)^2 = 1 + (10/2)^2 = 26$ is greater than $(\lambda_1/\phi_1)^2 = (4/2)^2 = 4$.

Figure 3(b) gives the difference in correlations between countries 1 and 3 with country 2 replaced by country 3 in (32). Here an alternative pattern to the one presented in Figure 3(a) emerges whereby the difference in correlations between countries 1 and 3 initially increases, then starts to decrease and eventually becomes negative for contagion levels of $\delta > 15$. For this case the inequality in (34) is now satisfied as $1+(\phi_3/\lambda_3)^2 = 1+(4/3)^2 = 2.778$ is less than $(\lambda_1/\phi_1)^2 = (4/2)^2 = 4$. This result is at odds with the usual strategy of testing for contagion based on identifying a significant increase in correlation during a crisis period. That is, the null hypothesis is usually taken as one-sided when testing for contagion based on correlation analysis; see also the discussion in Billio and Pelizzon (2003). However, this phenomenon is not uncommon in applied work when computing the correlation structure of asset returns during financial crises. What this result shows is that looking at simple correlations can be highly misleading when attempting to identify evidence of contagion. Even though for certain parameterization the correlation during the crisis period is less than it is for the noncrisis period, Figure 3(b) highlights that this is not inconsistent with the presence of contagion $(\delta > 0)$. From a testing point of view, this result also suggests that in testing for contagion based on correlation analysis, the power of the test statistic to detect contagion may not be monotonic over the parameter space. In fact, this test is not guaranteed to be unbiased, especially for one-sided null hypotheses.[8]

In contrast to the FR2 test, the FR1 and FR3 tests are based on comparing the crisis correlation with the total correlation. Defining the covariance and variance of the total period to be weighted averages of the crisis and noncrisis analogues, the expression in (32)

[8]The phenomenon that noncrisis correlations can exceed crisis correlations also suggests that the practice of choosing a crisis sample on the basis of the correlation in the crisis period exceeds the correlation in the non-crisis period is clearly inappropriate.

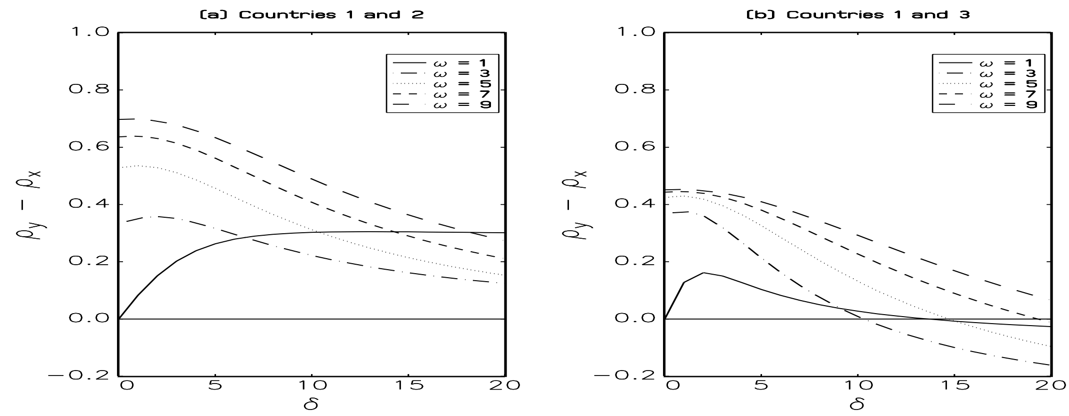

Figure 3: Differences in the crisis and pre-crisis correlations for alternative values of the contagion parameter ($\delta = \delta_2 = \delta_3$) and the structural break parameter (ω).

is now replaced by

$$\left(\rho_{y_{1,t},y_{2,t}} - \rho_{z_{1,t},z_{2,t}}\right) = \frac{\lambda_1\lambda_2 + \delta\phi_1^2}{\sqrt{\lambda_1^2+\phi_1^2}\sqrt{\lambda_2^2+\phi_2^2+\delta^2\phi_1^2}} - \frac{\lambda_1\lambda_2 T_x/T + \left(\lambda_1\lambda_2+\delta\phi_1^2\right)T_y/T}{\zeta}, \quad (35)$$

where

$$\zeta = \sqrt{\left(\lambda_1^2+\phi_1^2\right)T_x/T + \left(\lambda_1^2+\phi_1^2\right)T_y/T}\sqrt{\left(\lambda_2^2+\phi_2^2\right)T_x/T + \left(\lambda_2^2+\phi_2^2+\delta^2\phi_1^2\right)T_y/T}$$

and T_x and T_y are, as before, respectively the sample sizes of the noncrisis and crisis periods, and $T = T_x + T_y$ is the total sample period. Taking the limit as $\delta \to \infty$, yields

$$\lim_{\delta\to\infty}\left(\rho_{y_{1,t},y_{2,t}} - \rho_{z_{1,t},z_{2,t}}\right) = \frac{\phi_1}{\sqrt{\lambda_1^2+\phi_1^2}}\left(1 - \sqrt{\frac{T_y}{T}}\right) > 0.$$

Unlike the expression for $\lim_{\delta\to\infty}\left(\rho_{y_{1,t},y_{2,t}} - \rho_{x_{1,t},x_{2,t}}\right)$ in (33), the difference in the correlations is now guaranteed to be positive. This expression shows that as the relative proportion of the crisis period to the total period increases, the difference in the two correlations decreases. In the limit there would be no difference in the two correlations as the total and crisis periods would be the same. Figure 4 demonstrates the properties of (35) for the same parameterization used in Figure 3. In Figure 4(a) the difference in correlations is monotonic in δ initially, reaching a threshold for $\delta > 5$. In contrast, the pattern presented in Figure 4(b) is not monotonic for all values of δ, but is nonetheless still positive.

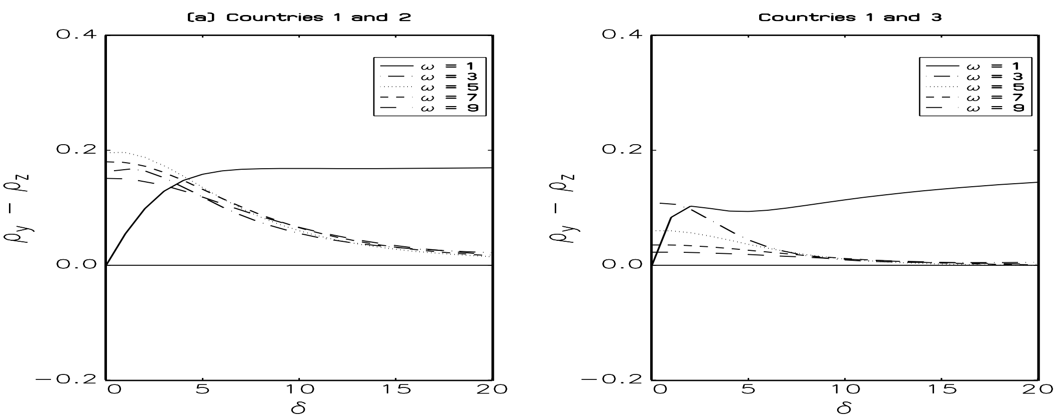

Figure 4: Differences in the crisis and total crisis correlations for alternative values of the contagion parameter ($\delta = \delta_2 = \delta_3$) and the structural break parameter (ω).

4.8.2 Spuriousness in Testing for Contagion

Contagion tests based on bivariate analysis such as the Forbes and Rigobon tests (FR1 and FR2), can potentially yield spurious contagious linkages between variables as a result of a common factor. This point is highlighted in Figure 5(a) in particular, which gives the difference in the crisis and noncrisis correlations between countries 2 and 3 for alternative values of the contagion parameter ($\delta = \delta_1 = \delta_2$) and the structural break parameter (ω). As the strength of contagion increases ($\delta > 0$), the correlation between the two asset returns increases. This increase in correlation is purely spurious as it arises from both variables being affected by a common factor, namely shocks from country 1 asset returns. A similar property is presented in Figure 5(b) which compares the crisis and total correlations between countries 2 and 3.

4.8.3 Structural Breaks

Loretan and English (2000) and Forbes and Rigobon (2002) emphasize the problems of structural breaks in testing for contagion. The problem is highlighted in Figures 3 to 5 for the case of a structural break in the common factor ($\omega > 1$). Increases in the strength of the structural break accentuates the differences in the noncrisis and crisis correlations. Figure 3(a) also shows that for greater levels of contagion, the switch in the magnitudes of the correlations in the two sample periods becomes even more marked.

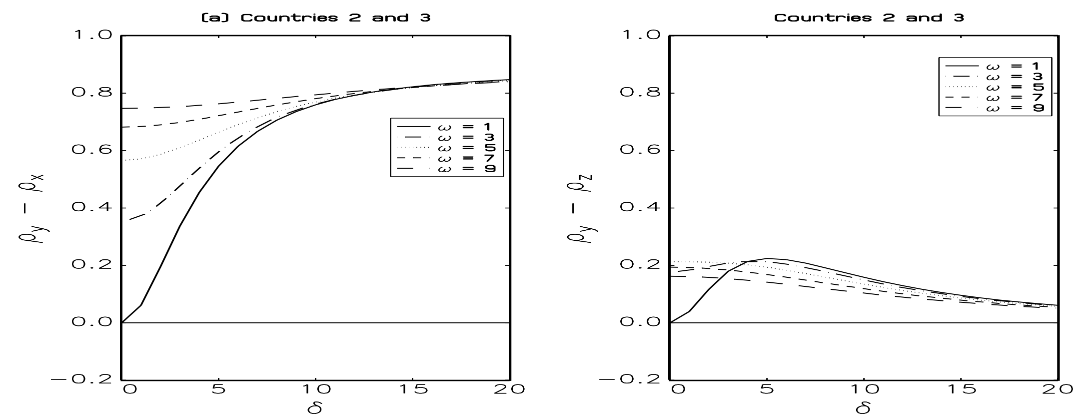

Figure 5: Differences in the crisis and pre-crisis correlations for alternative values of the contagion parameter ($\delta = \delta_2 = \delta_3$) and the structural break parameter (ω).

4.8.4 Weak Instruments

The Favero and Giavazzi test (FG) and the Pesaran and Pick test (PP1) both require identifying key parameters using instruments based on lagged variables. In applications using interest rates, as in the original Favero and Giavazzi application, the strong correlation in the data should potentially produce strong instruments. However, for applications using asset returns, as in the application by Walti (2003), the low levels of correlations in the data are unlikely to produce instruments with suitable properties. The effect of weak instruments can bias the IV estimator and result in asymmetric sampling distributions. In some cases the size of the IV bias could exceed the size of the OLS bias (Nelson and Startz (1990a, 1990b)). To try and isolate the effects of weak instruments, the Pesaran and Pick contagion tests (PP1 and PP2) are both reported.[9]

4.8.5 Information Loss from Filtering

The Favero and Giavazzi test (FG), the Pesaran and Pick tests (PP1 and PP2) and the Bae, Karolyi and Stulz test (BKS), all use a filter to identify large shocks. This contrasts with the Forbes and Rigobon class of tests (FR1, FR2, FR3 and FRM) which use all of the information in the sample to test for contagion. In general, the filtering methods represent a loss of information which can be expected to result in a loss of power. The extent of the loss in power is investigated in Monte Carlo experiments conducted below.

[9]Another strategy that can be adopted to gauge the size of weak instruments is to adopt the weak instrument tests suggested by Stock, Wright and Yogo (2002).

5 Finite Sample Properties

This section presents the finite sample properties of the eight contagion tests using a number of Monte Carlo experiments. The experiments are conducted to identify the size and power properties of the test statistics under various scenarios, including autocorrelation, structural breaks, GARCH conditional variance and unknown crisis periods.

5.1 Experimental Design

The DGP used in the Monte Carlo experiments is an extension of the DGP discussed in Section 3 by allowing for different types of structural breaks, autocorrelation in the common factor as well as time varying volatility based on a GARCH conditional variance. The model consists of three asset returns during a noncrisis period $(x_{1,t}, x_{2,t}, x_{3,t})$ and a crisis period $(y_{1,t}, y_{2,t}, y_{3,t})$. The crisis period is characterized by contagion from $y_{1,t}$ to both $y_{2,t}$, and $y_{3,t}$. The crisis period also allows for structural breaks in the common factor (w_t) and or the idiosyncratic factor of $y_{1,t}$.

Non-crisis Model

$$\begin{aligned} x_{1,t} &= 4w_t + 2u_{1,t} && (36) \\ x_{2,t} &= 2w_t + 10u_{2,t} && (37) \\ x_{3,t} &= 3w_t + 4u_{3,t}, && (38) \end{aligned}$$

where

$$\begin{aligned} w_t &= \rho w_{t-1} + u_{w,t} && (39) \\ u_{w,t} &\sim N(0, h_t) && (40) \\ h_t &= \omega^2 \left(1 - \alpha - \beta + \alpha u_{w,t-1}^2 + \beta h_{t-1}\right) && (41) \\ u_{i,t} &\sim N(0,1) \qquad i = 1, 2, 3. && (42) \end{aligned}$$

Crisis Model

$$\begin{aligned} y_{1,t} &= 4w_t + 2u_{1,t} && (43) \\ y_{2,t} &= 2w_t + 10u_{2,t} + \delta 2u_{1,t} && (44) \\ y_{3,t} &= 3w_t + 4u_{3,t} + \delta 2u_{1,t}, && (45) \end{aligned}$$

where

$$\begin{aligned} w_t &= \rho w_{t-1} + u_{w,t} && (46) \\ u_{w,t} &\sim N(0, h_t) && (47) \\ h_t &= \omega^2 \left(1 - \alpha - \beta + \alpha u_{w,t-1}^2 + \beta h_{t-1}\right) && (48) \\ u_{1,t} &\sim N\left(0, \kappa^2\right) && (49) \\ u_{i,t} &\sim N(0,1) \qquad i = 2, 3. && (50) \end{aligned}$$

The strength of contagion is controlled by the parameter δ in (44) and (45), with parameter values set at

$$\delta = \{0, 1, 2, 5, 10\}. \qquad (51)$$

A value of $\delta = 0$, represents no contagion and is used to examine the size properties of the test statistics in small samples when the asymptotic critical values are used. Values of $\delta > 0$, are used to examine the power properties of the contagion tests using size-adjusted critical values.

Two types of structural breaks are investigated. The first is a structural break in the common factor w_t. The parameter values chosen are

$$\omega = \{1, 5\}, \qquad (52)$$

where $\omega = 1$, represents no structural break in the common factor during the crisis period. The second is a structural break in the idiosyncratic factor of $y_{1,t}$, namely $u_{1,t}$. The parameter values are

$$\kappa = \{1, 5\}, \qquad (53)$$

where $\kappa = 1$, represents no structural break in the idiosyncratic factor during the crisis period.

The parameters α and β allow for a time varying volatility GARCH process common in financial market data. In the case where there is no conditional volatility

$$\alpha = \beta = 0.$$

For the conditional volatility case the parameters are set at

$$\alpha = 0.05, \beta = 0.90,$$

which are typical parameter estimates reported in empirical studies of financial returns (Engle, (2004)).

Eight Monte Carlo experiments are performed which are summarized in Table 5. For each experiment, six hypotheses are tested (Table 6) using eight alternative contagion tests (Table 4). The FR1, FR2 and FR3 tests are used to test the first four hypotheses in Table 6, but not the two joint hypotheses as these tests are bivariate by construction. The BKS test is not used to test the last joint hypothesis in Table 6 as this would involve stacking three sets of multivariate log likelihoods.

5.2 Computational Issues

All test statistics are based on a preliminary step consisting of extracting the residuals from estimating a trivariate VAR for $\{z_{1,t}, z_{2,t}, z_{3,t}\}$ with one lag over the full sample period. The Forbes and Rigobon set of univariate contagion tests (FR1, FR2, FR3) are evaluated by replacing $z_{i,t}$ in the pertinent formulae by the VAR residuals. This preliminary step is motivated by the empirical strategy adopted by Forbes and Rigobon (2002) who use it as a way of extracting out any common factors in the data.

The multivariate version of the Forbes and Rigobon test (FRM) is also obtained by replacing $z_{i,t}$ in (24) by the VAR residuals and estimating the set of equations by OLS. The FRM contagion test is evaluated using a Wald statistic.

The VAR residuals are used to compute the Favero and Giavazzi set of dummy variables defined in (26), which, in turn, are used in the structural model (27). This set of equations is estimated by IV which is asymptotically equivalent to FIML as the system of equations is just identified. Following the empirical work of Favero and Giavazzi, the contagion test (FG) is based on a likelihood ratio test.

The Pesaran and Pick contagion tests (PP1 and PP2) use the VAR residuals to compute the dummy variables in (28) which are used in the structural model in (29). The PP1 test is based on estimating the structural model by an IV estimator with the set of instruments consisting of a constant and all lagged variables $\{Const, z_{1,t-1}, z_{2,t-1}, z_{3,t-1}\}$. The PP2 test is obtained by ignoring potential simultaneity bias and replacing the IV estimator by the OLS estimator. Both contagion tests are evaluated using a Wald statistic.

The BKS contagion test (BKS) is based on using the VAR residuals to compute the exceedances and co-exceedances in (30). The BKS statistic is evaluated by estimating the

Table 5:

Summary of the Monte Carlo experiments.

Experiment	Type	Restrictions	Crisis Period
I	Strong Autocorrelation	$\omega = 1, \kappa = 1, \rho = 0.95,$ $\alpha = 0.0, \beta = 0.0$	Known
II	Weak Autocorrelation	$\omega = 1, \kappa = 1, \rho = 0.2,$ $\alpha = 0.0, \beta = 0.0$	Known
III	No Autocorrelation	$\omega = 1, \kappa = 1, \rho = 0.0,$ $\alpha = 0.0, \beta = 0.0$	Known
IV	Idiosyncratic structural break	$\omega = 1, \kappa = 5, \rho = 0.0,$ $\alpha = 0.0, \beta = 0.0$	Known
V	Common structural break	$\omega = 5, \kappa = 1, \rho = 0.0,$ $\alpha = 0.0, \beta = 0.0$	Known
VI	Common structural break GARCH conditional variance	$\omega = 5, \kappa = 1, \rho = 0.0,$ $\alpha = 0.05, \beta = 0.95$	Known
VII	Idiosyncratic structural break	$\omega = 1, \kappa = 5, \rho = 0.0,$ $\alpha = 0.0, \beta = 0.0$	Unknown
VIII	Common structural break	$\omega = 5, \kappa = 1, \rho = 0.0,$ $\alpha = 0.0, \beta = 0.0$	Unknown

Table 6:

Hypotheses tested in the Monte Carlo experiments.

Test	Hypothesis
$y_{1,t} \rightarrow y_{2,t}$	$H_0: \ \gamma_{2,1} = 0$
$y_{1,t} \rightarrow y_{3,t}$	$H_0: \ \gamma_{3,1} = 0$
$y_{2,t} \rightarrow y_{3,t}$	$H_0: \ \gamma_{3,2} = 0$
$y_{3,t} \rightarrow y_{2,t}$	$H_0: \ \gamma_{2,3} = 0$
$y_{1,t} \rightarrow y_{2,t}, y_{3,t}$	$H_0: \ \gamma_{2,1} = \gamma_{3,1} = 0$
$y_{1,t} \rightarrow y_{2,t}, y_{3,t}$ $y_{2,t} \rightarrow y_{3,t}$ $y_{3,t} \rightarrow y_{2,t}$	$H_0: \ \gamma_{2,1} = \gamma_{3,1} = \gamma_{3,2} = \gamma_{2,3} = 0$

system of equations in (31) by MLE. The BFGS algorithm in GAUSS is used to maximize the likelihood, with the maximum number of iterations set at 200. This is a gradient algorithm with the derivatives computed numerically.[10]

In the case where the crisis period is known, the Forbes and Rigobon tests (FR1, FR2, FR3, FRM) use all of the information in the crisis period to compute crisis correlations and define dummy variables as is needed in the case of the FRM test. The threshold dummy variables defined in (26) for the FG test are identified as the largest 5% of values during the crisis period. The same rule is adopted in (28) for the PP1 and PP2 tests. The exceedances and co-exceedances in (30) for the BKS tests are identified as the largest 10% of values during the crisis period.

In the experiments where the crisis period is unknown the rules for each of the tests are chosen as follows. The crisis period for the Forbes and Rigobon tests is selected as points in time corresponding to the largest observation in all returns. To avoid choosing a break point to close to the boundaries of the data, the first and last ten observations are excluded from identifying the start of the crisis period. The FG crisis period is determined by the

[10] The use of a VAR to extract out the dependence structure in the data is problematic in the case of the BKS approach which uses a multivariate logit model based on the assumption that the dichotomous variables are independent, although not identical. To introduce dependence directly into the model through either the mean (autocorrelation) or the variance (GARCH), would yield a more complicated likelihood function which would make estitmation more complicated.

rule given in (26). The PP crisis period is determined by the rule given in(28). Finally, the BKS crisis period is identified by the exceedances and co-exceedances as defined in (30).

In computing the size and power of the FG test in the case of the crisis sample period being unknown, the critical value will change in each draw of the Monte Carlo experiment as the number of outliers and hence the number of dummy variables that are being tested will change. To compute the size and power in this case the approach adopted is to perform the test using the asymptotic critical value from a chi-squared distribution where the degrees of freedom equalled the number of dummy variables being tested. The size and power are then simply computed as the sample mean of the number of times the null is rejected. In the experiments where the sample size is known (Experiments I to VI) the asymptotic distribution is simply χ^2_5 as the number of extreme observations is always set at 5 by construction.

All Monte Carlo experiments are conducted in Gauss 6.0, with each experiment based on $10,000$ replications. The normal random numbers are generated using the GAUSS procedure $RNDN$, with a seed equal to 123457. The noncrisis sample size is set at $T_x = 100$, while the sample for the crisis period is chosen as $T_y = 50$.

5.3 Size

The finite sample results of the size of the eight contagion tests across the eight experiments, are presented in Table 7 (Experiments I to IV) and Table 8 (Experiments V to VIII). Figures 6 and 7 gives the full sampling distributions across the eight experiments for each test in the case of testing the hypothesis of contagion form $y_{1,t}$ to $y_{2,t}$.[11] The results are based on simulating the model under the null hypothesis of no contagion by setting $\delta = 0$. The sizes are based on the 5% (one-sided) asymptotic normal critical values in the case of the FR1, FR2 and FR3 tests, while the sizes of the remaining test statistics are based on the 5% asymptotic chi-squared critical values.

The sampling distributions of the FR1, FR2 and FR3 test statistics under the null hypothesis of no contagion are given in the top nine panels of Figure 6 for testing $y_{1,t} \rightarrow y_{2,t}$. For comparison, the asymptotic distribution, $N(0,1)$, is also presented. The part of the right tail of the sampling distribution of the FRM test statistic is given in the bottom three panels of Figure 6 to test the same hypothesis, while Figure 7 contains the tails of the sam-

[11] The reported sampling distributions are based on a normal kernel with the bandwidth computed using a normal reference rule in the case of FR1, FR2 and FR3. For the remaining tests, the sampling distribution is based on a logarithmic transformation kernel to ensure postivity of the sampling distribution.

pling distribution of the FG, PP1, PP2 and BKS test statistics. The asymptotic distribution also included in the figures of the FRM, PP1, PP2 and BKS statistics for comparison is the χ_1^2 distribution, whereas for the FG statistic it is the χ_5^2 distribution.[12]

The FR1 statistic is consistently undersized (less than 0.050) for all experiments with the test not rejecting the null of no contagion often enough. The problem is especially acute for Experiments IV to VIII where there is a structural break in one of the factors. This feature is highlighted in Figure 6 for the case of testing contagion from $y_{1,t}$ to $y_{2,t}$.

The poor size properties of the FR1 statistic in the case of Experiments I to III, stem from using an asymptotic variance based on incorrectly assuming independence between the crisis and noncrisis (total in this case) samples. This point is highlighted in Figure 6 for Experiments I to III, where the FR2 and FR3 test statistics generate sampling distributions that closely approximate the asymptotic distribution. Table 7 shows that both of these tests consistently yield sizes close to 0.05 for the first three experiments for all types of hypotheses tested.

The FR1 and FR3 tests also exhibit low sizes in the presence of structural breaks. The FR2 test also yields low sizes in the case of idiosyncratic structural breaks (Experiments IV and VII), but reasonable sizes when structural breaks occur in the common factor. This is true for when the crisis period is known (Experiment IV), there is GARCH conditional volatility (Experiment VI) and the crisis period is unknown (Experiment VIII). This result suggests that by splitting up the sample into two distinct sub-periods and not defining the noncrisis period as the total period, provides a more robust test of contagion.

The FRM test also demonstrates good size properties for a range of experiments and hypotheses tested, especially for the first three experiments. The bottom panels of Figure 6 show that the tails of the sampling distribution for these three experiments are close to the right tail of the asymptotic sampling distribution (χ_1^2). The size of FRM generally tends to be slightly inflated compared to the sizes of FR2 and FR3, which reflects an efficiency loss in conducting the FRM test in small samples as a result of the inclusion of variables that are not part of the DGP. In contrast to the FR1, FR2 and FR3 tests, the FRM statistic appears to be relatively robust to structural breaks, common or idiosyncratic, for most of the hypotheses investigated, but not all. An exception is in testing for contagion from $y_{1,t}$ to $y_{3,t}$ in Experiment IV, where the size of the test is inflated in excess of 0.95.

[12] The sampling distribution of the FG statistic is χ_5^2, in the case where the crisis period is known (Experiements I to VI), while in the case where the crisis period is unknown it is a mixture of χ^2 distributions.

The FG test is consistently oversized for all experiments and across all hypotheses tested. The top three panels of Figure 7 highlight the fatness in the tails of the sampling distribution in the case of testing the first hypothesis, namely contagion from $y_{1,t}$ to $y_{2,t}$. Comparing the sizes of the FG statistic across Experiments I to III, shows that the sizes increase as the level of autocorrelation is reduced. This highlights the weak instrument problems with this test as a reduction in autocorrelation leads to a deterioration in the quality of the instruments yielding sampling distributions with fatter tails. In the extreme case where there is no autocorrelation (Experiment III), the simultaneous equations model underlying the FG test is underidentified and the instruments are now irrelevant.

A feature of the size properties of the PP1 statistic is that it is consistently undersized. The thinness in the tails of the distribution are highlighted in Figure 7 for the case of the first hypothesis being tested, with the tails of all experiments lying below the tail of the asymptotic distribution. This result is in stark contrast to the size properties of the FG test which is consistently oversized, but in agreement with the FR1 statistic which is also consistently undersized in all experiments. The poor size properties of the test reflects the loss of information in modelling contagion through the specification of dummy variables. This loss of information would appear to be particularly severe as this effect is able to dominate the weak instrument effect which could have the opposite result of inflating the size of the statistic. This point is highlighted by looking at the size properties of the PP2 test statistic. As this statistic uses more information than PP1, its size properties should be marginally better. This is true in the case of Experiment I, but for all other experiments the sizes are practically the same as they are for PP1.

The size properties of the BKS statistic are similar to the FG statistic as it also yields relatively inflated sizes across most experiments and across most hypotheses. An exception is in Experiment VII where the sizes are reasonable, ranging between 0.039 to 0.095 for all of the four single tests of contagion.

5.4 Power

Tables 9 to 16 give the probability of finding contagion, adjusted by size, for each of the eight tests for increasing intensity levels of contagion, $\delta = 1, 2, 5, 10$, across the eight experiments. As contagion is assumed to run from $y_{1,t}$ to both $y_{2,t}$ and $y_{3,t}$ during the crisis period, the power of the test should increase monotonically as δ increases for the first two hypotheses,

$y_{1,t} \rightarrow y_{2,t}$ and $y_{1,t} \rightarrow y_{3,t}$. For the third and fourth hypotheses, $y_{2,t} \rightarrow y_{3,t}$ and $y_{3,t} \rightarrow y_{2,t}$, the power should be equal to the size adjusted value of the test, namely 0.05. Although the size of the FR3 test is closer to the asymptotic distribution under the null than FR1, the power functions of the FR1 and FR3 tests are identical as the use of the size adjusted critical values merely rescales the relevant distributions.

5.4.1 Autocorrelation: Experiments I to III

Tables 9 to 11 give the size adjusted power functions for the alternative contagion tests for high (Table 9), low (Table 10) and zero (Table 11) levels of autocorrelation. In testing for contagion from $y_{1,t}$ to $y_{2,t}$ and $y_{3,t}$, the FR1 and FR3 tests both exhibit monotonic power functions with the power of the test increasing for increasing levels of autocorrelation. The FR2 test exhibits slightly better power in testing for contagion from $y_{1,t}$ to $y_{2,t}$, but its power function is non-monotonic in testing for contagion from $y_{1,t}$ to $y_{3,t}$. This latter result reflects the properties of the correlation function presented in Figure 3(b), where it is revealed that the correlation between two series can decrease for relatively high levels of contagion and even become negative. The FR2 and FRM tests exhibit similar power functions with the latter test yielding marginally inferior power properties.

Tables 9 to 11 reveal that the remaining tests (FG, PP1, PP2 and BKS) all exhibit very low and flat power functions in testing for contagion from $y_{1,t}$ to $y_{2,t}$ and $y_{3,t}$. To try and unravel the poor performance of these tests a comparison of the PP1 and PP2 reveals approximately a 100% improvement in power in testing for contagion from $y_{1,t}$ to $y_{2,t}$ and $y_{3,t}$, for high levels of autocorrelation (Table 9). This result is interpreted as a problem of weak instruments with the PP1 test with the bias arising from using weak instruments (PP1) exceeding the simultaneity bias (PP2). However, as the strength of autocorrelation falls (Tables 10 and 11), the improvement in power between PP1 and PP2 tapers off with the powers of the two tests tending to converge for the extreme case where there is no autocorrelation (Table 11). One way to understand this result is to compare the power properties of the PP2 test with the power of FRM test, which also does not correct for potential simultaneity bias. The main difference in the two tests is how contagion is identified. With the PP2 test only "big" shocks are included whereas with the FRM test all shocks are included in testing for contagion. The reduction in power of the PP2 test then represents the information loss of using a filter that excludes important sample information.

While a comparison of the power properties of the FR1, FR2 and FR3 in testing the first two hypotheses reveals that the FR2 exhibits slightly better power, this comes at a cost in testing for contagion between $y_{2,t}$ and $y_{3,t}$, where the test incorrectly identifies contagion with the probability of detection increasing as the strength of contagion increases. The power functions of the FR1 and FR3 tests corresponding to these two hypotheses exhibit non-monotonic functions with the power first increasing and then decreasing for higher levels of contagion.

5.4.2 Structural Breaks: Experiments IV to VI

The size adjusted power functions of alternative tests of contagion are given in Table 12 (idiosyncractic structural break), Table 12 (common structural break) and Table 14 (common structural break with conditional volatility).

Table 12 shows that the bivariate FR tests (FR1, FR2 and FR3) all exhibit high power in testing contagion in the presence of idiosyncractic structural breaks from $y_{1,t}$ to $y_{2,t}$ and $y_{3,t}$, with power quickly approaching unity for even "moderate" levels of contagion in most cases. In contrast, all remaining contagion tests exhibit low power with at best, very flat power functions across the contagion spectrum.

Table 12 also shows that the power in testing for contagion between $y_{2,t}$ and $y_{3,t}$ using the FR1 and FR2 tests incorrectly yield a probability in excess of 0.05 of finding contagion for "low" levels of contagion, but this probability falls as the strength of contagion increases and approaches the correct level of 0.05 for "large" values of contagion. In contrast, the FR2 and FRM tests incorrectly find contagion between $y_{2,t}$ and $y_{3,t}$ and that the power function quickly approaches unity for the FRM test, especially in testing for contagion from $y_{2,t}$ to $y_{3,t}$. This result suggests that by choosing the noncrisis period as the total period (FR1 and FR3) yields a more robust test of contagion especially in the case where there are structural breaks in the idiosyncratic factor. Of course, the Forbes and Rigobon test is designed to capture idiosyncratic structural breaks by using the adjusted correlation coefficient given in (17). In regards to the other tests of contagion, especially FG, PP1 and PP2, whilst these tests yield powers close to the theoretical value of 0.05, this is simply the result of these tests exhibiting very little power for any hypothesis test at any level of contagion in the first place.

Table 13 shows that all tests have low power in the presence of a structural break in the common factor. An exception is the FRM test which exhibits increasing power in testing for contagion from $y_{1,t}$ and $y_{2,t}$, however this test also incorrectly has increasing power in testing for contagion between $y_{2,t}$ and $y_{3,t}$. By allowing for a GARCH conditional variance in the common factor (Table 14) does not change the power properties of the tests in the presence of a structural break in the common factor.

5.4.3 Crisis Period Unknown: Experiments VII to VIII

The power properties of the contagion tests in the presence of idiosyncratic and common structural breaks are given in Tables 15 and 16 respectively, when the crisis period is unknown. Comparing these results with the respective power properties in Tables 12 and 13 where the crisis period is known, not surprisingly reveals that there is a slight loss of power with the FR1, FR2 and FR3 tests. As in the case where the crisis period is known, the other tests still exhibit low power.

5.4.4 Joint Tests of Contagion: Experiments I to VIII

The last two columns of Tables 9 to 16 give the power functions of two joint tests of contagion for selected contagion tests. The first test is a test that there is contagion from $y_{1,t}$ to both $y_{2,t}$ and $y_{3,t}$, while the second test is a joint of overall contagion in all directions. The FRM and BKS tests perform the best. The FRM test exhibits highest power for low and moderate levels of contagion, but its power function begins to fall for high levels of contagion as a result of the effects of the falling power detected in testing for contagion from $y_{1,t}$ to $y_{3,t}$, discussed previously. At this point in the power comparison, the BKS yields relatively higher power.

6 Conclusions

This paper has investigated the finite sample properties of a range of tests of contagion commonly employed to detect propagation mechanisms during financial crises. The tests investigated included the Forbes and Rigobon adjusted correlation test, the Favero and Giavazzi outlier test, the Pesaran and Pick threshold test, and the Bae, Karolyi and Stulz co-exceedance test. Some extensions of these tests were also proposed to adjust for size distortions and weak instruments. The experiments conducted allowed for autocorrelation in the mean and the variance, alternative types of structural breaks and situations where the timing of the crisis period may either be known or unknown.

The results showed that the Forbes and Rigobon test is undersized. Two variants of this test were proposed which were shown to correct the size distortion. The first adjustment consisted of not defining the noncrisis period as the total period, while the second approach involved deriving a new asymptotic standard error which took into account the lack of dependence between the crisis and noncrisis periods as adopted in the original version of the Forbes and Rigobon test. As with the original version of the Forbes and Rigobon test, the Pesaran and Pick test was also undersized, whereas the Favero and Giavazzi test and the Bae, Karolyi and Stulz test are oversized.

For certain parameterization some of the test statistics exhibited non-monotonic power functions. This was shown to be less of a problem in the case of the Forbes and Rigobon tests which defined the noncrisis period as the total sample. The Favero and Giavazzi test and the Pesaran and Pick test exhibited low power which was the result of a combination of weak instrument problems and, more importantly, a loss of information from the types of filters employed in these tests to identify large shocks during crisis periods. The Bae, Karolyi and Stulz also exhibited low power arising from the adoption of filters designed to identify large shocks.

Most tests perform badly in terms of size and power in the presence of structural breaks. One exception is the Forbes and Rigobon test where the noncrisis and crisis periods are independent, which has reasonable size properties when there are idiosyncratic structural breaks. Another exception is a multivariate version of the Forbes and Rigobon test proposed by Dungey, Fry, González-Hermosillo and Martin, which tends to have reasonable size properties for not only idiosyncratic structural breaks, but also common structural breaks.

The theoretical results of the paper have a number of important implications for undertaking tests of contagion. First, the property that the Forbes and Rigobon test is undersized is consistent with much of the empirical literature where empirical studies find little evidence of contagion when using this test. A quick fix is either to define the crisis and noncrisis periods as disjoint samples, or if the noncrisis period is defined as the total sample period then the correct asymptotic standard error, as derived in this paper, can be used.

Second, the Pesaran and Pick test is also undersized. This suggests that this test is also a conservative test thereby it will tend not to find evidence of contagion when it does exist.

Third, and in contrast to the Forbes and Rigobon test and the Pesaran and Pick test, the Favero and Giavazzi test and the Bae, Karolyi and Stulz test, are oversized. Again this is consistent with the empirical findings where it is common to find evidence of contagion when using these tests. This suggests that much of the empirical evidence of contagion using these tests are potentially spurious, arising simply from adopting a test where the effective probability of a Type I error is in excess of the usual level.

Fourth, when using the Favero and Giavazzi test and the Pesaran and Pick test, care should be given to potential weak instrument problems. This is especially true for applications conducted on financial returns and the instruments chosen are the lagged returns. One solution is to use a test that is robust to weak instruments (see Stock, Wright and Yogo (2002) for a survey).

Fifth, care must be given to modelling structural breaks in testing for contagion otherwise the tests can be badly biased. One solution is to use the multivariate version of the Forbes and Rigobon test which was found to have the best robustness properties to the presence of structural breaks.

A Proof of the Asymptotic Variance of the Adjusted Forbes and Rigobon Test

In this appendix, the standard error for the adjusted Forbes and Rigobon test is derived. Let r_y be the sample correlation for the second sample period (T_y), and r represent the correlation for the total sample period $(T = T_x + T_y)$ where T_x is the first sample period. To construct a test of the difference in the second sub-sample and total correlations, it is necessary to find

$$Var(r_y - r) = Var(r_y) + Var(r) - 2Cov(r_y, r) . \tag{54}$$

Focussing on the last term

$$Cov(r_y, r) = Cov\left(\frac{m_{11}}{\sqrt{m_{20}m_{02}}}, \frac{n_{11}}{\sqrt{n_{20}n_{02}}}\right), \tag{55}$$

where m_{11} (n_{11}) is the sample covariance in the second (total) period, and $m_{20} = m_{02}$ $(n_{20} = n_{02})$ is the sample standard deviation in the second (total) period. Define

$$v_1 = m_{11}, v_2 = m_{20}, v_3 = m_{02}, v_4 = n_{11}, v_5 = n_{20}, v_6 = n_{02},$$

where the respective population values are θ_i, and the functions

$$g(v) = \frac{v_1}{\sqrt{v_2 v_3}}, h(v) = \frac{v_4}{\sqrt{v_5 v_6}},$$

where $v = \{v_1, v_2, \cdots, v\}$.

For a first order approximation (Kendall and Stuart, (1969, Vol.1, p.232))

$$Cov(r_y, r) = \sum_{i=1}^{6} g_i'(\theta) h_i'(\theta) Var(v_i) + \sum_{i=1}^{6} \sum_{\substack{j=1 \\ i \neq j}}^{6} g_i'(\theta) h_i'(\theta) Cov(v_i, v_j), \tag{56}$$

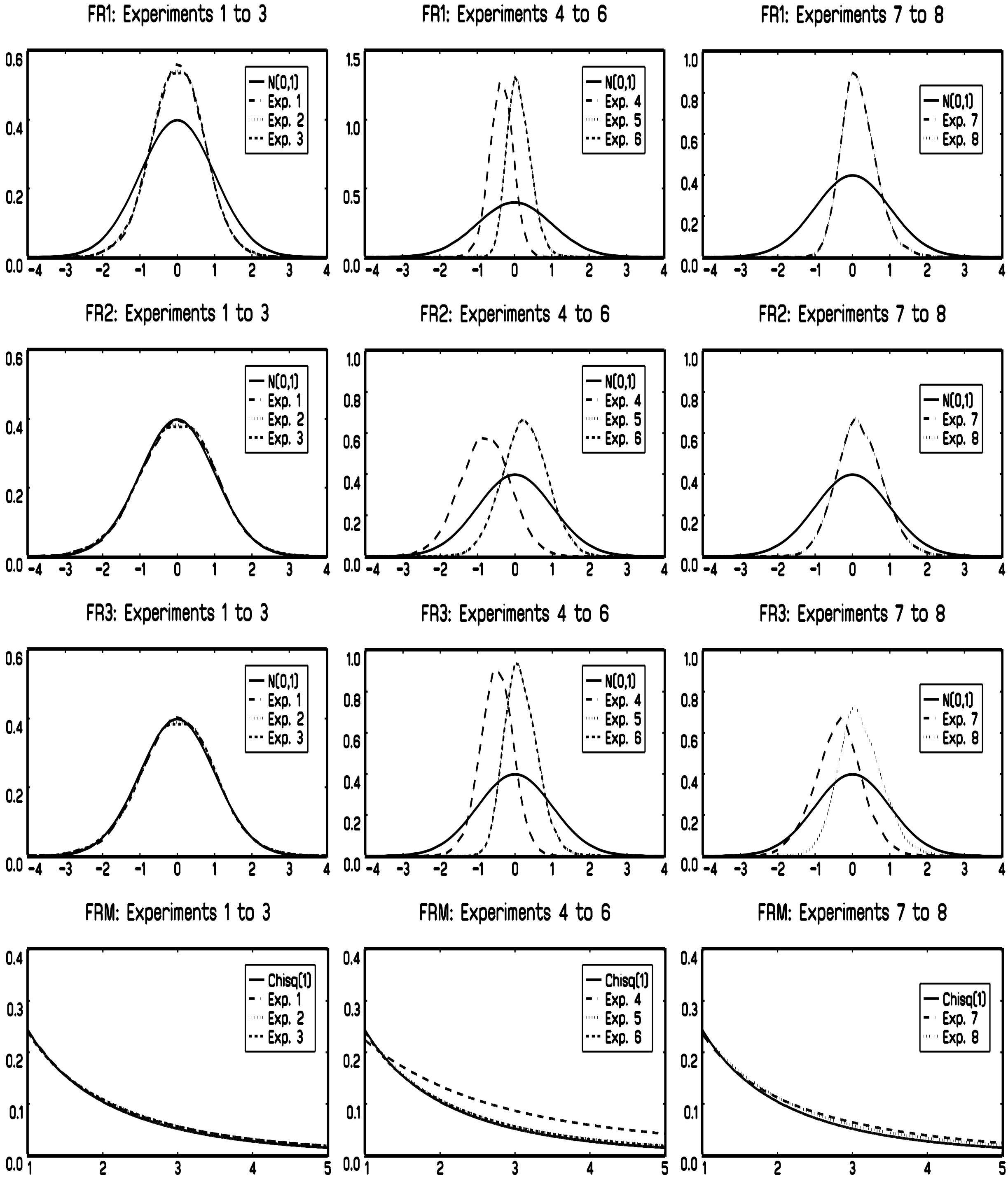

Figure 6: Sampling distributions of FR1, FR2, FR3, FRM under the null of no contagion to test the hypothesis $y_{1,t} \rightarrow y_{2,t}$, for experiments 1 to 6. The range of the sampling distribution for FRM is restricted to focus on the upper tail properties.

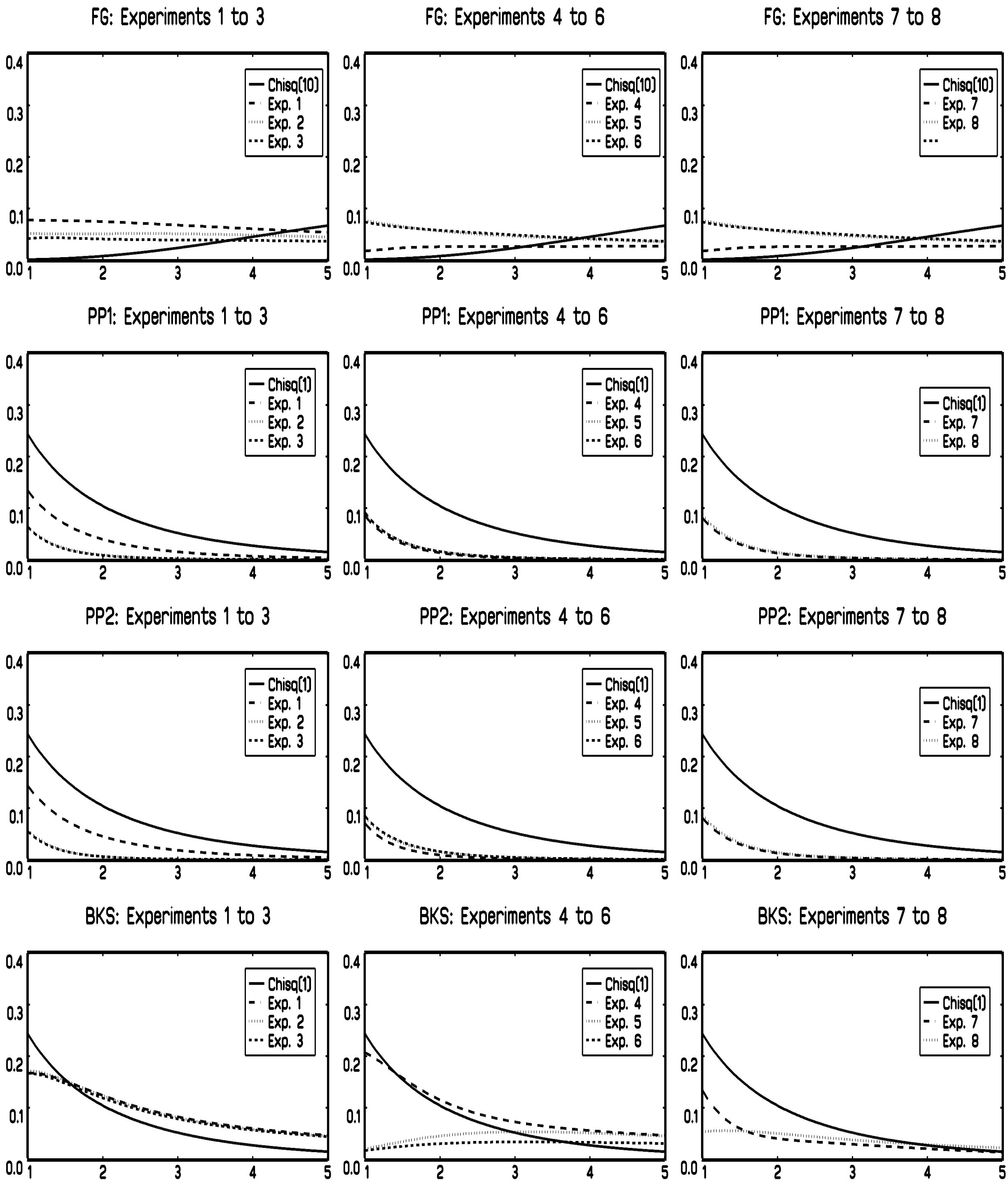

Figure 7: Sampling distributions of FG, PP1, PP2, BKS under the null of no contagion to test the hypothesis $y_{1,t} \rightarrow y_{2,t}$, for experiments 1 to 6. The range of the sampling distributions are restricted to focus on the upper tail properties.

Table 7:

Size properties of alternative contagion tests based on an asymptotic size of 5%: Experiments I, II, III and IV ($\delta = 0$).

Experiment	Test	Hypothesis					
		$1 \to 2$	$1 \to 3$	$2 \to 3$	$3 \to 2$	$1 \to 2,3$	Total
I	FR1	0.012	0.009	0.013	0.013	n.a.	n.a.
	FR2	0.049	0.039	0.051	0.051	n.a.	n.a.
	FR3	0.050	0.045	0.051	0.052	n.a.	n.a.
	FRM	0.060	0.097	0.057	0.057	0.092	0.104
	FG	0.596	0.669	0.896	0.806	0.646	0.973
	PP1	0.006	0.006	0.010	0.008	0.003	0.026
	PP2	0.009	0.013	0.005	0.005	0.006	0.027
	BKS	0.113	0.256	0.277	0.107	0.191	n.a.
II	FR1	0.013	0.009	0.012	0.012	n.a.	n.a.
	FR2	0.049	0.043	0.051	0.052	n.a.	n.a.
	FR3	0.050	0.049	0.052	0.053	n.a.	n.a.
	FRM	0.059	0.094	0.056	0.058	0.091	0.106
	FG	0.703	0.748	0.841	0.783	0.780	0.979
	PP1	0.000	0.000	0.000	0.000	0.000	0.000
	PP2	0.000	0.000	0.000	0.000	0.000	0.000
	BKS	0.112	0.261	0.280	0.108	0.175	n.a.
III	FR1	0.013	0.009	0.012	0.013	n.a.	n.a.
	FR2	0.051	0.043	0.052	0.052	n.a.	n.a.
	FR3	0.052	0.049	0.054	0.054	n.a.	n.a.
	FRM	0.058	0.095	0.055	0.058	0.090	0.107
	FG	0.769	0.804	0.838	0.776	0.878	0.989
	PP1	0.000	0.000	0.000	0.000	0.000	0.000
	PP2	0.000	0.000	0.000	0.000	0.000	0.000
	BKS	0.112	0.253	0.275	0.119	0.181	n.a.
IV	FR1	0.000	0.000	0.012	0.013	n.a.	n.a.
	FR2	0.000	0.000	0.052	0.052	n.a.	n.a.
	FR3	0.000	0.000	0.054	0.054	n.a.	n.a.
	FRM	0.183	0.985	0.101	0.076	0.979	0.953
	FG	0.855	0.876	0.882	0.859	0.945	0.997
	PP1	0.000	0.000	0.000	0.000	0.000	0.000
	PP2	0.000	0.000	0.000	0.000	0.000	0.000
	BKS	0.131	0.154	0.165	0.129	0.084	n.a.

Table 8:

Size properties of alternative contagion tests based on an asymptotic size of 5%: Experiments V, VI, VII and VIII ($\delta = 0$).

Experiment	Test	Hypothesis					
		$1 \to 2$	$1 \to 3$	$2 \to 3$	$3 \to 2$	$1 \to 2, 3$	Total
V	FR1	0.000	0.003	0.227	0.002	n.a.	n.a.
	FR2	0.010	0.080	0.939	0.178	n.a.	n.a.
	FR3	0.003	0.025	0.548	0.024	n.a.	n.a.
	FRM	0.065	0.381	0.066	0.065	0.336	0.320
	FG	0.527	0.604	0.913	0.548	0.649	0.983
	PP1	0.001	0.000	0.000	0.000	0.000	0.000
	PP2	0.001	0.004	0.001	0.000	0.001	0.002
	BKS	0.367	0.712	0.158	0.315	0.406	n.a.
VI	FR1	0.000	0.003	0.221	0.002	n.a.	n.a.
	FR2	0.011	0.082	0.921	0.175	n.a.	n.a.
	FR3	0.003	0.026	0.537	0.024	n.a.	n.a.
	FRM	0.060	0.099	0.055	0.057	0.096	0.109
	FG	0.537	0.609	0.911	0.553	0.659	0.983
	PP1	0.000	0.000	0.000	0.001	0.000	0.000
	PP2	0.001	0.003	0.001	0.000	0.001	0.002
	BKS	0.367	0.712	0.158	0.315	0.406	n.a.
VII	FR1	0.001	0.000	0.015	0.014	n.a.	n.a.
	FR2	0.002	0.000	0.051	0.052	n.a.	n.a.
	FR3	0.002	0.000	0.051	0.053	n.a.	n.a.
	FRM	0.084	0.463	0.064	0.055	0.425	0.370
	FG	0.756	0.774	0.235	0.230	0.861	0.920
	PP1	0.000	0.000	0.000	0.000	0.000	0.000
	PP2	0.000	0.000	0.000	0.000	0.000	0.000
	BKS	0.039	0.095	0.051	0.046	0.015	n.a.
VIII	FR1	0.007	0.015	0.173	0.015	n.a.	n.a.
	FR2	0.017	0.037	0.420	0.055	n.a.	n.a.
	FR3	0.014	0.030	0.287	0.030	n.a.	n.a.
	FRM	0.070	0.129	0.071	0.069	0.180	0.161
	FG	0.520	0.596	0.226	0.300	0.635	0.826
	PP1	0.000	0.000	0.000	0.000	0.000	0.000
	PP2	0.000	0.025	0.000	0.001	0.011	0.015
	BKS	0.179	0.934	0.086	0.108	0.378	n.a.

Table 9:

Experiment I: Power properties of alternative contagion tests, based on size adjusted critical values of 5% ($\delta > 0$).

Intensity	Test	Hypothesis					
		$1 \to 2$	$1 \to 3$	$2 \to 3$	$3 \to 2$	$1 \to 2, 3$	Total
$\delta = 1.0$	FR1	0.115	0.288	0.080	0.060	n.a.	n.a.
	FR2	0.119	0.295	0.091	0.067	n.a.	n.a.
	FR3	0.115	0.288	0.080	0.060	n.a.	n.a.
	FRM	0.070	0.187	0.049	0.054	0.181	0.123
	FG	0.035	0.033	0.041	0.046	0.032	0.038
	PP1	0.042	0.053	0.044	0.041	0.050	0.045
	PP2	0.051	0.066	0.042	0.040	0.059	0.057
	BKS	0.064	0.109	0.049	0.061	0.104	n.a.
$\delta = 2.0$	FR1	0.209	0.378	0.184	0.087	n.a.	n.a.
	FR2	0.220	0.394	0.249	0.125	n.a.	n.a.
	FR3	0.209	0.378	0.184	0.087	n.a.	n.a.
	FRM	0.061	0.439	0.104	0.089	0.384	0.411
	FG	0.034	0.028	0.040	0.047	0.026	0.040
	PP1	0.053	0.074	0.051	0.053	0.067	0.067
	PP2	0.066	0.103	0.048	0.042	0.095	0.085
	BKS	0.082	0.148	0.064	0.089	0.162	n.a.
$\delta = 5.0$	FR1	0.434	0.328	0.404	0.075	n.a.	n.a.
	FR2	0.481	0.189	0.810	0.280	n.a.	n.a.
	FR3	0.434	0.328	0.404	0.075	n.a.	n.a.
	FRM	0.102	0.481	0.983	0.720	0.416	0.997
	FG	0.051	0.040	0.041	0.065	0.042	0.050
	PP1	0.074	0.103	0.068	0.072	0.097	0.113
	PP2	0.103	0.152	0.062	0.056	0.150	0.147
	BKS	0.144	0.098	0.349	0.377	0.327	n.a.
$\delta = 10.0$	FR1	0.573	0.463	0.218	0.046	n.a.	n.a.
	FR2	0.629	0.076	0.923	0.394	n.a.	n.a.
	FR3	0.573	0.463	0.218	0.046	n.a.	n.a.
	FRM	0.168	0.197	1.000	0.938	0.179	1.000
	FG	0.078	0.068	0.042	0.060	0.063	0.056
	PP1	0.068	0.089	0.068	0.062	0.079	0.093
	PP2	0.119	0.137	0.056	0.049	0.135	0.121
	BKS	0.111	0.011	0.667	0.624	0.334	n.a.

Table 10:

Experiment II: Power properties of alternative contagion tests, based on size adjusted critical values of 5% ($\delta > 0$).

Intensity	Test	Hypothesis					
		$1 \rightarrow 2$	$1 \rightarrow 3$	$2 \rightarrow 3$	$3 \rightarrow 2$	$1 \rightarrow 2, 3$	Total
$\delta = 1.0$	FR1	0.120	0.307	0.080	0.067	n.a.	n.a.
	FR2	0.124	0.315	0.089	0.074	n.a.	n.a.
	FR3	0.120	0.307	0.080	0.067	n.a.	n.a.
	FRM	0.075	0.203	0.050	0.052	0.193	0.134
	FG	0.047	0.045	0.052	0.048	0.046	0.045
	PP1	0.048	0.042	0.047	0.047	0.045	0.044
	PP2	0.052	0.049	0.046	0.046	0.052	0.048
	BKS	0.063	0.101	0.045	0.066	0.105	n.a.
$\delta = 2.0$	FR1	0.222	0.438	0.203	0.116	n.a.	n.a.
	FR2	0.237	0.462	0.264	0.160	n.a.	n.a.
	FR3	0.222	0.438	0.203	0.116	n.a.	n.a.
	FRM	0.066	0.491	0.103	0.087	0.438	0.463
	FG	0.041	0.040	0.051	0.048	0.042	0.047
	PP1	0.052	0.046	0.044	0.052	0.045	0.044
	PP2	0.056	0.053	0.046	0.045	0.059	0.058
	BKS	0.080	0.146	0.059	0.089	0.162	n.a.
$\delta = 5.0$	FR1	0.445	0.318	0.536	0.114	n.a.	n.a.
	FR2	0.512	0.240	0.898	0.429	n.a.	n.a.
	FR3	0.445	0.318	0.536	0.114	n.a.	n.a.
	FRM	0.096	0.518	0.981	0.716	0.462	0.998
	FG	0.052	0.050	0.050	0.061	0.049	0.054
	PP1	0.062	0.049	0.045	0.066	0.055	0.060
	PP2	0.069	0.062	0.062	0.068	0.075	0.112
	BKS	0.155	0.122	0.329	0.365	0.354	n.a.
$\delta = 10.0$	FR1	0.511	0.339	0.315	0.046	n.a.	n.a.
	FR2	0.629	0.088	0.995	0.554	n.a.	n.a.
	FR3	0.511	0.339	0.315	0.046	n.a.	n.a.
	FRM	0.159	0.205	1.000	0.937	0.183	1.000
	FG	0.072	0.070	0.045	0.056	0.058	0.051
	PP1	0.063	0.045	0.047	0.069	0.054	0.071
	PP2	0.078	0.053	0.089	0.100	0.077	0.163
	BKS	0.108	0.017	0.639	0.611	0.350	n.a.

Table 11:

Experiment III: Power properties of alternative contagion tests, based on size adjusted critical values of 5% ($\delta > 0$).

Intensity	Test	Hypothesis					
		$1 \to 2$	$1 \to 3$	$2 \to 3$	$3 \to 2$	$1 \to 2, 3$	Total
$\delta = 1.0$	FR1	0.115	0.306	0.079	0.067	n.a.	n.a.
	FR2	0.119	0.313	0.088	0.074	n.a.	n.a.
	FR3	0.115	0.306	0.079	0.067	n.a	n.a.
	FRM	0.072	0.204	0.050	0.053	0.193	0.134
	FG	0.043	0.047	0.054	0.048	0.045	0.045
	PP1	0.048	0.046	0.051	0.046	0.044	0.045
	PP2	0.046	0.050	0.046	0.045	0.052	0.048
	BKS	0.070	0.110	0.046	0.059	0.102	n.a.
$\delta = 2.0$	FR1	0.215	0.441	0.199	0.115	n.a.	n.a.
	FR2	0.228	0.462	0.262	0.161	n.a.	n.a.
	FR3	0.215	0.441	0.199	0.115	n.a.	n.a.
	FRM	0.063	0.494	0.103	0.087	0.436	0.465
	FG	0.038	0.043	0.056	0.049	0.041	0.041
	PP1	0.051	0.046	0.046	0.050	0.051	0.047
	PP2	0.051	0.054	0.045	0.050	0.055	0.058
	BKS	0.087	0.152	0.062	0.087	0.165	n.a.
$\delta = 5.0$	FR1	0.429	0.316	0.529	0.113	n.a.	n.a.
	FR2	0.503	0.240	0.900	0.429	n.a.	n.a.
	FR3	0.429	0.316	0.529	0.113	n.a.	n.a.
	FRM	0.094	0.522	0.981	0.717	0.460	0.998
	FG	0.050	0.056	0.052	0.057	0.048	0.046
	PP1	0.062	0.051	0.054	0.068	0.059	0.072
	PP2	0.066	0.066	0.073	0.076	0.079	0.121
	BKS	0.151	0.140	0.332	0.353	0.349	n.a.
$\delta = 10.0$	FR1	0.496	0.333	0.308	0.047	n.a.	n.a.
	FR2	0.620	0.089	0.905	0.555	n.a.	n.a.
	FR3	0.496	0.333	0.308	0.047	n.a.	n.a.
	FRM	0.159	0.205	1.000	0.937	0.182	1.000
	FG	0.070	0.065	0.043	0.056	0.049	0.048
	PP1	0.065	0.053	0.058	0.073	0.065	0.091
	PP2	0.072	0.066	0.105	0.102	0.082	0.184
	BKS	0.113	0.023	0.649	0.612	0.348	n.a.

Table 12:

Experiment IV: Power properties of alternative contagion tests, based on size adjusted critical values of 5% ($\delta > 0$).

Intensity	Test	Hypothesis					
		$1 \to 2$	$1 \to 3$	$2 \to 3$	$3 \to 2$	$1 \to 2, 3$	Total
$\delta = 1.0$	FR1	0.848	1.000	0.529	0.113	n.a.	n.a.
	FR2	0.925	1.000	0.900	0.429	n.a.	n.a.
	FR3	0.848	1.000	0.529	0.113	n.a.	n.a.
	FRM	0.057	0.007	0.029	0.041	0.019	0.037
	FG	0.021	0.033	0.058	0.031	0.030	0.034
	PP1	0.042	0.039	0.033	0.044	0.039	0.039
	PP2	0.052	0.076	0.052	0.049	0.075	0.088
	BKS	0.169	0.741	0.084	0.102	0.464	n.a.
$\delta = 2.0$	FR1	0.970	1.000	0.308	0.047	n.a.	n.a.
	FR2	1.000	1.000	0.995	0.555	n.a.	n.a.
	FR3	0.970	1.000	0.308	0.047	n.a.	n.a
	FRM	0.017	0.425	0.535	0.360	0.451	1.000
	FG	0.023	0.028	0.051	0.027	0.026	0.031
	PP1	0.034	0.030	0.029	0.037	0.034	0.039
	PP2	0.053	0.061	0.054	0.055	0.069	0.122
	BKS	0.135	0.513	0.141	0.179	0.500	n.a.
$\delta = 5.0$	FR1	0.984	1.000	0.055	0.028	n.a.	n.a.
	FR2	1.000	1.000	0.993	0.525	n.a.	n.a.
	FR3	0.984	1.000	0.055	0.028	n.a.	n.a.
	FRM	0.027	0.022	1.000	0.822	0.027	1.000
	FG	0.051	0.046	0.039	0.031	0.029	0.031
	PP1	0.031	0.026	0.027	0.028	0.031	0.033
	PP2	0.039	0.042	0.052	0.050	0.050	0.117
	BKS	0.038	0.066	0.219	0.252	0.121	n.a.
$\delta = 10.0$	FR1	0.985	1.000	0.031	0.026	n.a.	n.a.
	FR2	1.000	1.000	0.930	0.396	n.a.	n.a.
	FR3	0.985	1.000	0.031	0.026	n.a.	n.a.
	FRM	0.024	0.000	1.000	0.698	0.000	1.000
	FG	0.084	0.080	0.034	0.034	0.032	0.032
	PP1	0.021	0.021	0.020	0.019	0.020	0.023
	PP2	0.028	0.029	0.036	0.034	0.036	0.078
	BKS	0.003	0.004	0.165	0.175	0.011	n.a.

Table 13:

Experiment V: Power properties of alternative contagion tests, based on size adjusted critical values of 5% ($\delta > 0$).

Intensity	Test	Hypothesis					
		$1 \to 2$	$1 \to 3$	$2 \to 3$	$3 \to 2$	$1 \to 2, 3$	Total
$\delta = 1.0$	FR1	0.054	0.124	0.049	0.050	n.a.	n.a.
	FR2	0.055	0.105	0.052	0.051	n.a.	n.a.
	FR3	0.054	0.124	0.049	0.050	n.a.	n.a.
	FRM	0.053	0.068	0.048	0.051	0.068	0.059
	FG	0.050	0.045	0.058	0.052	0.049	0.057
	PP1	0.044	0.046	0.047	0.044	0.043	0.045
	PP2	0.048	0.046	0.049	0.047	0.045	0.045
	BKS	0.046	0.053	0.043	0.044	0.049	n.a.
$\delta = 2.0$	FR1	0.032	0.014	0.046	0.050	n.a.	n.a.
	FR2	0.050	0.023	0.055	0.057	n.a.	n.a.
	FR3	0.032	0.014	0.046	0.050	n.a.	n.a.
	FRM	0.054	0.040	0.125	0.101	0.031	0.079
	FG	0.056	0.053	0.054	0.050	0.052	0.064
	PP1	0.055	0.054	0.060	0.056	0.054	0.059
	PP2	0.056	0.049	0.057	0.057	0.050	0.058
	BKS	0.055	0.044	0.077	0.074	0.061	n.a.
$\delta = 5.0$	FR1	0.001	0.000	0.024	0.041	n.a.	n.a.
	FR2	0.021	0.000	0.079	0.090	n.a.	n.a.
	FR3	0.001	0.000	0.024	0.041	n.a.	n.a.
	FRM	0.354	0.005	0.994	0.769	0.081	0.927
	FG	0.101	0.080	0.050	0.067	0.078	0.074
	PP1	0.088	0.083	0.095	0.087	0.088	0.114
	PP2	0.084	0.063	0.106	0.094	0.074	0.108
	BKS	0.049	0.009	0.464	0.416	0.083	n.a.
$\delta = 10.0$	FR1	0.000	0.000	0.004	0.024	n.a.	n.a.
	FR2	0.005	0.000	0.117	0.148	n.a.	n.a.
	FR3	0.000	0.000	0.004	0.024	n.a.	n.a.
	FRM	0.493	0.030	1.000	0.945	0.357	1.000
	FG	0.157	0.100	0.044	0.077	0.097	0.083
	PP1	0.097	0.089	0.110	0.100	0.097	0.146
	PP2	0.087	0.061	0.129	0.116	0.073	0.135
	BKS	0.016	0.000	0.768	0.584	0.025	n.a.

Table 14:

Experiment VI: Power properties of alternative contagion tests, based on size adjusted critical values of 5% ($\delta > 0$).

Intensity	Test	Hypothesis					
		$1 \to 2$	$1 \to 3$	$2 \to 3$	$3 \to 2$	$1 \to 2, 3$	Total
$\delta = 1.0$	FR1	0.055	0.126	0.049	0.050	n.a.	n.a.
	FR2	0.055	0.102	0.052	0.052	n.a.	n.a.
	FR3	0.055	0.126	0.049	0.050	n.a.	n.a.
	FRM	0.075	0.195	0.051	0.053	0.189	0.131
	FG	0.046	0.047	0.051	0.045	0.052	0.051
	PP1	0.044	0.044	0.046	0.047	0.045	0.048
	PP2	0.045	0.045	0.048	0.047	0.042	0.045
	BKS	0.046	0.053	0.043	0.044	0.049	n.a.
$\delta = 2.0$	FR1	0.034	0.014	0.046	0.049	n.a.	n.a.
	FR2	0.051	0.024	0.057	0.058	n.a.	n.a.
	FR3	0.034	0.014	0.046	0.049	n.a.	n.a.
	FRM	0.066	0.491	0.103	0.087	0.441	0.466
	FG	0.055	0.054	0.054	0.048	0.054	0.051
	PP1	0.055	0.054	0.056	0.060	0.058	0.059
	PP2	0.057	0.053	0.057	0.058	0.054	0.057
	BKS	0.055	0.044	0.077	0.074	0.061	n.a.
$\delta = 5.0$	FR1	0.001	0.000	0.025	0.041	n.a.	n.a.
	FR2	0.021	0.000	0.080	0.092	n.a.	n.a.
	FR3	0.001	0.000	0.025	0.041	n.a.	n.a.
	FRM	0.093	0.529	0.979	0.713	0.478	0.998
	FG	0.093	0.088	0.045	0.063	0.089	0.065
	PP1	0.093	0.081	0.094	0.094	0.095	0.122
	PP2	0.084	0.069	0.104	0.093	0.074	0.112
	BKS	0.049	0.009	0.464	0.416	0.083	n.a.
$\delta = 10.0$	FR1	0.000	0.000	0.004	0.024	n.a.	n.a.
	FR2	0.005	0.000	0.121	0.150	n.a.	n.a.
	FR3	0.000	0.000	0.004	0.024	n.a.	n.a.
	FRM	0.156	0.218	1.000	0.937	0.201	1.000
	FG	0.148	0.114	0.040	0.072	0.105	0.072
	PP1	0.099	0.092	0.108	0.107	0.100	0.152
	PP2	0.089	0.064	0.125	0.113	0.075	0.139
	BKS	0.016	0.000	0.768	0.584	0.025	n.a.

Table 15:

Experiment VII: Power properties of alternative contagion tests, based on size adjusted critical values of 5% ($\delta > 0$).

Intensity	Test	Hypothesis					
		$1 \to 2$	$1 \to 3$	$2 \to 3$	$3 \to 2$	$1 \to 2, 3$	Total
$\delta = 1.0$	FR1	0.447	0.904	0.400	0.129	n.a.	n.a.
	FR2	0.487	0.934	0.408	0.129	n.a.	n.a.
	FR3	0.460	0.904	0.287	0.069	n.a.	n.a.
	FRM	0.048	0.006	0.053	0.065	0.025	0.022
	FG	0.021	0.031	0.077	0.089	0.033	0.045
	PP1	0.044	0.040	0.046	0.062	0.045	0.056
	PP2	0.072	0.151	0.061	0.087	0.149	0.190
	BKS	0.284	0.979	0.117	0.123	0.473	n.a.
$\delta = 2.0$	FR1	0.687	0.957	0.261	0.079	n.a.	n.a.
	FR2	0.756	0.976	0.374	0.096	n.a.	n.a.
	FR3	0.693	0.957	0.170	0.041	n.a.	n.a
	FRM	0.059	0.057	0.093	0.093	0.129	0.325
	FG	0.010	0.016	0.119	0.175	0.014	0.059
	PP1	0.042	0.035	0.041	0.060	0.041	0.055
	PP2	0.067	0.132	0.095	0.123	0.139	0.321
	BKS	0.377	0.848	0.566	0.587	0.808	n.a.
$\delta = 5.0$	FR1	0.770	0.955	0.088	0.059	n.a.	n.a.
	FR2	0.838	0.977	0.181	0.059	n.a.	n.a.
	FR3	0.774	0.955	0.049	0.033	n.a.	n.a.
	FRM	0.079	0.008	0.217	0.141	0.111	0.552
	FG	0.014	0.016	0.134	0.211	0.009	0.056
	PP1	0.037	0.036	0.043	0.048	0.042	0.054
	PP2	0.069	0.086	0.174	0.170	0.102	0.431
	BKS	0.351	0.276	0.701	0.704	0.976	n.a.
$\delta = 10.0$	FR1	0.773	0.957	0.062	0.055	n.a.	n.a.
	FR2	0.839	0.978	0.097	0.042	n.a.	n.a.
	FR3	0.776	0.957	0.035	0.031	n.a.	n.a.
	FRM	0.083	0.003	0.191	0.151	0.099	0.595
	FG	0.031	0.031	0.122	0.172	0.011	0.051
	PP1	0.029	0.031	0.035	0.037	0.033	0.040
	PP2	0.056	0.063	0.156	0.156	0.082	0.364
	BKS	0.156	0.075	0.486	0.487	0.992	n.a.

Table 16:

Experiment VIII: Power properties of alternative contagion tests, based on size adjusted critical values of 5% ($\delta > 0$).

Intensity	Test	Hypothesis					
		$1 \to 2$	$1 \to 3$	$2 \to 3$	$3 \to 2$	$1 \to 2, 3$	Total
$\delta = 1.0$	FR1	0.052	0.083	0.047	0.050	n.a.	n.a.
	FR2	0.052	0.089	0.048	0.050	n.a.	n.a.
	FR3	0.053	0.091	0.047	0.049	n.a.	n.a.
	FRM	0.048	0.047	0.044	0.052	0.047	0.047
	FG	0.051	0.050	0.050	0.047	0.051	0.046
	PP1	0.046	0.048	0.044	0.047	0.044	0.044
	PP2	0.047	0.047	0.052	0.050	0.048	0.048
	BKS	0.052	0.045	0.043	0.043	0.051	n.a.
$\delta = 2.0$	FR1	0.045	0.025	0.043	0.051	n.a.	n.a.
	FR2	0.049	0.026	0.046	0.050	n.a.	n.a.
	FR3	0.045	0.023	0.044	0.049	n.a.	n.a.
	FRM	0.069	0.058	0.085	0.064	0.068	0.081
	FG	0.054	0.052	0.054	0.058	0.053	0.052
	PP1	0.053	0.057	0.056	0.056	0.055	0.055
	PP2	0.051	0.053	0.057	0.056	0.053	0.057
	BKS	0.051	0.039	0.083	0.085	0.078	n.a.
$\delta = 5.0$	FR1	0.018	0.001	0.028	0.050	n.a.	n.a.
	FR2	0.022	0.001	0.045	0.050	n.a.	n.a.
	FR3	0.016	0.001	0.029	0.048	n.a.	n.a.
	FRM	0.170	0.302	0.624	0.208	0.377	0.556
	FG	0.076	0.064	0.065	0.103	0.064	0.071
	PP1	0.087	0.084	0.074	0.082	0.095	0.104
	PP2	0.060	0.039	0.106	0.102	0.043	0.083
	BKS	0.037	0.001	0.546	0.571	0.175	n.a.
$\delta = 10.0$	FR1	0.007	0.001	0.008	0.040	n.a.	n.a.
	FR2	0.013	0.001	0.036	0.046	n.a.	n.a.
	FR3	0.006	0.000	0.007	0.036	n.a.	n.a.
	FRM	0.109	0.313	0.786	0.165	0.410	0.795
	FG	0.113	0.074	0.080	0.143	0.067	0.090
	PP1	0.095	0.094	0.095	0.098	0.106	0.141
	PP2	0.064	0.021	0.187	0.153	0.025	0.125
	BKS	0.052	0.000	0.944	0.944	0.121	n.a.

where a $'$ denotes the first derivative. Using the expressions for $g(x)$ and $h(x)$

$$\begin{aligned}
g_1'(v) &= \frac{1}{\sqrt{v_2 v_3}} \\
g_2'(v) &= -\frac{1}{2} v_1 v_2^{-3/2} v_3^{-1/2} \\
g_3'(v) &= -\frac{1}{2} v_1 v_2^{-1/2} v_3^{-3/2} \\
g_4'(v) &= g_5'(v) = g_6'(v) = 0 \\
h_1'(v) &= h_2'(v) = h_3'(v) = 0 \\
h_4'(v) &= \frac{1}{\sqrt{v_5 v_6}} \\
h_5'(v) &= -\frac{1}{2} v_4 v_5^{-3/2} v_6^{-1/2} \\
h_6'(v) &= -\frac{1}{2} v_4 v_5^{-1/2} v_6^{-3/2}.
\end{aligned}$$

Substituting into the expression for the covariance in (56) and evaluating at the population value θ, gives

$$\begin{aligned}
Cov(r_y, r) &= (\theta_2\theta_3)^{-1/2}(\theta_5\theta_6)^{-1/2} Cov(v_1, v_4) \\
&\quad -\frac{1}{2}\theta_2^{-1/2}\theta_3^{-1/2}\theta_4\theta_5^{-3/2}\theta_6^{-1/2} Cov(v_1, v_5) \\
&\quad -\frac{1}{2}\theta_2^{-1/2}\theta_3^{-1/2}\theta_4\theta_5^{-3/2}\theta_6^{-1/2} Cov(v_1, v_6) \\
&\quad -\frac{1}{2}\theta_1\theta_2^{-3/2}\theta_3^{-1/2}\theta_5^{-1/2}\theta_6^{-1/2} Cov(v_2, v_4) \\
&\quad +\frac{1}{4}\theta_1\theta_2^{-3/2}\theta_3^{-1/2}\theta_4\theta_5^{-3/2}\theta_6^{-1/2} Cov(v_2, v_5) \\
&\quad +\frac{1}{4}\theta_1\theta_2^{-3/2}\theta_3^{-1/2}\theta_4\theta_5^{-1/2}\theta_6^{-3/2} Cov(v_2, v_5) \\
&\quad -\frac{1}{2}\theta_1\theta_2^{-1/2}\theta_3^{-3/2}\theta_5^{-1/2}\theta_6^{-1/2} Cov(v_3, v_4) \\
&\quad +\frac{1}{4}\theta_1\theta_2^{-1/2}\theta_3^{-3/2}\theta_4\theta_5^{-3/2}\theta_6^{-1/2} Cov(v_3, v_5) \\
&\quad +\frac{1}{4}\theta_1^2\theta_2^{-1/2}\theta_3^{-3/2}\theta_4\theta_5^{-1/2}\theta_6^{-3/2} Cov(v_3, v_6).
\end{aligned}$$

Using the result that under the the null, the distribution is standardized bivariate normal, then

$$\theta_1 = \theta_4 = \rho, \theta_2 = \theta_3 = \theta_5 = \theta_6 = 1,$$

and the expression for $Cov\left(r_y, r\right)$ simplifies to

$$\begin{aligned} Cov\left(r_y, r\right) &= Cov\left(v_1, v_4\right) - \frac{\rho}{2}\left(Cov\left(v_1, v_5\right) + Cov\left(v_1, v_6\right) + Cov\left(v_2, v_4\right) + Cov\left(v_3, v_4\right)\right) \\ &\quad + \frac{\rho^2}{4}\left(Cov\left(v_2, v_5\right) + Cov\left(v_2, v_6\right) + Cov\left(v_3, v_5\right) + Cov\left(v_3, v_6\right)\right). \qquad (57) \end{aligned}$$

This expression simplifies even further for the case of contemporaneous independence ($\rho = 0$)

$$Cov\left(r_y, r\right) = Cov\left(v_1, v_4\right). \qquad (58)$$

Now consider the sample covariances

$$\begin{aligned} v_1 &= m_{11} = \frac{1}{T_y}\sum_{t=1}^{T_y}\left(y_{1,t} - \overline{y}_1\right)\left(y_{2,t} - \overline{y}_2\right) \\ v_4 &= n_{11} = \frac{1}{T}\sum_{t=1}^{T}\left(z_{1,t} - \overline{z}_1\right)\left(z_{2,t} - \overline{z}_2\right). \qquad (59) \end{aligned}$$

It is helpful to rewrite the covariance for the total sample period as

$$\begin{aligned} n_{11} &= \frac{1}{T}\sum_{t=1}^{T}\left(z_{1,t} - \overline{z}_1\right)\left(z_{2,t} - \overline{z}_2\right) \\ &= \frac{1}{T}\left(\sum_{t=1}^{T_x}\left(x_{1,t} - \overline{x}_1\right)\left(x_{2,t} - \overline{x}_2\right) + \sum_{t=T_x+1}^{T}\left(y_{1,t} - \overline{y}_1\right)\left(y_{2,t} - \overline{y}_2\right)\right) \\ &= \frac{1}{T}\left(\sum_{t=1}^{T_x}\left(x_{1,t} - \overline{x}_1\right)\left(x_{2,t} - \overline{x}_2\right)\right) + \frac{T_y}{T}m_{11}, \end{aligned}$$

so

$$Cov\left(v_1, v_4\right) = Cov\left(n_{11}, \frac{1}{T}\left(\sum_{t=1}^{T_x}\left(x_{1,t} - \overline{x}_1\right)\left(x_{2,t} - \overline{x}_2\right)\right) + \frac{T_y}{T}m_{11}\right). \qquad (60)$$

Because of independence, this expression immediately reduces to

$$Cov\left(v_1, v_4\right) = Cov\left(n_{11}, \frac{T_y}{T}m_{11}\right) = \frac{T_y}{T}Var\left(m_{11}\right) \qquad (61)$$

An expression for $Var\left(m_{11}\right)$ is obtained using the property of the standard errors of bivariate moments (Kendall and Stuart, (1969, Vol.1, p.232))

$$Var\left(m_{11}\right) = \frac{1}{T_y}\left(\mu_{2,2} - \mu_{1,1}^2 + \mu_{2,0}\mu_{0,1}^2 + \mu_{0,2}\mu_{1,0}^2 + 2\mu_{1,1}\mu_{0,1}\mu_{1,0} - 2\mu_{2,1}\mu_{0,1} - 2\mu_{1,2}\mu_{1,0}\right), \qquad (62)$$

which reduces to

$$\begin{aligned} Var\left(m_{11}\right) &= \frac{1}{T_y}\left(\mu_{2,2}-\mu_{1,1}^2\right) \\ &= \frac{1}{T_y}\left(1+2\rho^2-\rho^2\right) \end{aligned} \tag{63}$$

as $\mu_{0,1} = \mu_{1,0} = 0$ because the random variables have zero mean, and for the standardized bivariate normal distribution $\mu_{2,2} = 1 + 2\rho^2, \mu_{1,1}^2 = \rho^2$. For the case of contemporaneous independence ($\rho = 0$) then $Var\left(m_{11}\right) = T_y^{-1}$ and from (61)

$$Cov\left(r_y, r\right) = \frac{T_y}{T}\frac{1}{T_y} = \frac{1}{T}. \tag{64}$$

Finally, using the result that the distribution is bivariate normal and contemporaneously independent, $Var\left(r_y\right) = T_y^{-1}$ and $Var\left(r\right) = T^{-1}$, gives the key result

$$\begin{aligned} Var\left(r_y - r\right) &= Var\left(r_y\right) + Var\left(r\right) - 2Cov\left(r_y, r\right) \\ &= \frac{1}{T_y} + \frac{1}{T} - 2\frac{1}{T} \\ &= \frac{1}{T_y} - \frac{1}{T}. \end{aligned} \tag{65}$$

References

[1] Bae, K.H, Karolyi, G.A. and Stulz, R.M. (2003), "A New Approach to Measuring Financial Contagion", *Review of Financial Studies*, 16(3), 717-763.

[2] Baig, T. and Goldfajn, I. (1999), "Financial Market Contagion in the Asian Crisis", *IMF Staff Papers,* 46(2), 167-195.

[3] Baur, D. and Schulze, N. (2002), "Coexceedances in Financial Markets - A Quantile Regression Analysis of Contagion", University of Tuebingen Discussion Paper 253.

[4] Bekaert, G., Harvey, C.R. and Ng, A. (2005), "Market Integration and Contagion", *Journal of Business*, 78(1), part 2, forthcoming.

[5] Billio, M. and Pelizzon, L. (2003), "Contagion and Interdependence in Stock Markets: Have They Been Misdiagnosed?", *Journal of Economics and Business*, 55 405-426.

[6] Ciccarelli, M. and Rebucci, A. (2004), "Measuring Contagion and Interdependence with a Bayesian Time-Varying Coefficient Model: The Chilean FX Market During the Argentine Crisis", mimeo.

[7] Corsetti, G., Pericoli, M. and Sbracia, M. (2002), "Some Contagion, Some Interdependence': More Pitfalls in Testing for Contagion", mimeo University of Rome III.

[8] Dasgupta, A. (2000), "Regionality Revisited: An Examination of the Direction of Spread of Currency Crises", mimeo, LSE.

[9] Dornbusch, R., Park, Y.C. and Claessens, S. (2000), "Contagion: Understanding How It Spreads", *The World Bank Research Observer*, 15(2), 177-97.

[10] Dungey, M., Fry, R., González-Hermisillo, B. and Martin, V.L. (2002), "The Transmission of Contagion in Developed and Developing International Bond Markets", in Committee on the Global Financial System (ed), *Risk Measurement and Systemic Risk*, Proceedings of the Third Joint Central Bank Research Conference, 61-74 .

[11] Dungey, M., Fry, R., Gonzalez-Hermosillo, B. and Martin, V.L. (2005a), "A Comparison of Alternative Tests of Contagion with Applications", in M. Dungey and D. Tambakis (eds) *Identifying International Financial Contagion: Progress and Challenges*, Oxford University Press, 60-85.

[12] Dungey, M., Fry, R., González-Hermosillo, B. and Martin, V.L. (2005b), "Empirical Modelling of Contagion: A Review of Methodologies", *Quantitative Finance*, 5(1), 9-24.

[13] Dungey, M. and Martin, V.L. (2004), "A Multifactor Model of Exchange Rates with Unanticipated Shocks: Measuring Contagion in the East Asian Currency Crisis", *Journal of Emerging Markets Finance*, 3, 305-330.

[14] Dungey, M., Fry, R., González-Hermosillo, B. and Martin, V.L. (2005c) LTCM *Journal of Financial Stability*, forthcoming.

[15] Dungey, M., Martin, V.L. and Pagan, A.R. (2000), "A Multivariate Latent Factor Decomposition of International Bond Yield Spreads", *Journal of Applied Econometrics*, 15, 697-715.

[16] Eichengreen, B., Rose, A.K. and Wyplosz, C. (1995), "Exchange Market Mayhem: The Antecedents and Aftermath of Speculative Attacks", *Economic Policy*, 21, 249-312.

[17] Eichengreen, B., Rose, A.K. and Wyplosz, C. (1996), "Contagious Currency Crises", NBER Working Paper, 5681.

[18] Engle, R.F., (2004), "Risk and Volatility: Econometric Models and Financial Practice, *American Economic Review*, 94, 405-420.

[19] Favero, C.A. and Giavazzi, F. (2002), "Is the International Propagation of Financial Shocks Non-linear? Evidence From the ERM", *Journal of International Economics*, 57 (1), 231-46.

[20] Forbes, K. and Rigobon, R. (2002), "No Contagion, Only Interdependence: Measuring Stock Market Co-movements", *Journal of Finance*, 57 (5), 2223-61.

[21] Glick, R. and Rose, A.K. (1999), "Contagion and Trade: Why are Currency Crises Regional?", *Journal of International Money and Finance*, 18(4), 603-17.

[22] Hahn, J. and Hausman, J. (2002), "A New Specification Test for the Validity of Instrumental Variables", *Econometrica*, 70, 163-189.

[23] Jeanne, O., and Masson, P. (2000), "Currency Crises, Sunspots and Markov-Switching Regimes", *Journal of International Economics,* 50(2), 327-350.

[24] Kaminsky, G.L. and Reinhart, C.M. (2000), "On Crises, Contagion and Confusion", *Journal of International Economics,* 51(1), 145-168.

[25] Kaminsky, G.L. and Reinhart, C.M. (2001), "Financial Markets in Times of Stress", NBER Working Paper, 8569.

[26] Kendall, M. and Stuart, A. (1969), *The Advanced Theory of Statistics*, Vol.1, Charles Griffin and Co., London.

[27] Kendall, M. and Stuart, A. (1973), *The Advanced Theory of Statistics*, Vol.2, Charles Griffin and Co., London.

[28] Loretan, M. and English, W. (2000), "Evaluating "Correlation Breakdowns" During Periods of Market Volatility", Board of Governors of the Federal Reserve System, International Finance Discussion Paper No. 658.

[29] Lowell, J., Neu, C.R. and Tong, D. (1998), "Financial Crises and Contagion in Emerging Market Countries", Monograph, RAND.

[30] Mody, A, and Taylor, M.P. (2003), "Common Vulnerabilities", CEPR Discussion Paper 3759.

[31] Nelson, C.R. and Startz, R. (1990a), "Some Further Results on the Exact Small Sample Properties of the Instrumental Variables Estimator", *Econometrica*, 58, 967-976.

[32] Nelson, C.R. and Startz, R. (1990b), "The Distribution of the Instrumental Variables Estimator and Its t-Ratio When the Instrument is a Poor One", *Journal of Business*, 63, S125-S140.

[33] Pesaran, H. and Pick, A. (2003), "Econometric Issues in the Analysis of Contagion", CESifo Working Paper 1176.

[34] Pericoli, M. and Sbracia, M. (2003), "A Primer on Financial Contagion", *Journal of Economic Surveys*, 17(4), 571-608.

[35] Pritsker, M. (2002), "Large Investors and Liquidity: A Review of the Literature", in Committee on the Global Financial System (ed), *Risk Measurement and Systemic Risk*, Proceedings of the Third Joint Central Bank Research Conference.

[36] Rigobon, R. (2003), "On the Measurement of the International Propogation of Shocks: Is the Transmission Stable?", *Journal of International Economics*, 61, 261-283.

[37] Stock, J.H., Wright, J.H. and Yogo, M. (2002), "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments", *Journal of Business and Economic Statistics*, 20(4), 518-529.

[38] Walti, S. (2003), "Testing for Contagion in International Financial Markets: Which Way to Go?", The Graduate Institute of International Studies, Geneva, HEI Working Paper No: 04/2003.