

Measuring Contagion with a Bayesian, Time-Varying Coefficient Model*

Matteo Ciccarelli[†]
European Central Bank
and
Alessandro Rebucci[‡]
International Monetary Fund

First Draft: October 2002

This Draft: 05 June 2003

Abstract

To measure contagion empirically, we propose using a Bayesian time-varying coefficient model estimated with Markov Chain Monte Carlo methods. The proposed measure works in the joint presence of heteroskedasticity and omitted variables and does not require knowledge of the timing of the crisis. It distinguishes contagion not only from interdependence but also from structural breaks. It can be used to investigate positive as well as negative contagion. The proposed measure appears to work well using both simulated and actual data.

Key Words: Contagion, Gibbs sampling, Heteroskedasticity, Omitted variable bias, Time-varying coefficient models.

Jel Classification Codes: C11, C15, F41, F42, G15.

*We are thankful to Tam Bayoumi, Pablo Garcia, Philipp Hartmann, Jérôme Henry, Eduardo Ley, Jens Nystedt, Steve Phillips, Roberto Rigobon, and seminar participants at the Central Bank of Chile, ECARES-Free University of Bruxelles, Ente Einaudi-Bank of Italy, IGIER-Bocconi University, University of Alicante, European Central Bank, International Monetary Fund, and the 2002 LAMES in Sao Paulo for useful comments and discussions. Ciccarelli's research was done in part while he was at the University of Alicante and benefitted from financial support by IVIE, DGIMCYT (BEC2002-03097), and Generalitat Valenciana (CTIDIB/2002/175). The views expressed in this paper are exclusively those of the authors and not those of the European Central Bank or the International Monetary Fund. Remaining errors are of the authors.

[†]European Central Bank, DG Research, Kaiserstrasse 29, D-60311, Frankfurt am Main. E-mail: matteo.ciccarelli@ecb.int.

[‡]Corresponding author, International Monetary Fund, Research Department, 700 19th St., N.W., Washington, D.C. 20431, USA. Phone: +1 (202) 623 4837. Fax +1 (202) 589 4837. E-mail: arebucci@imf.org.

1 Introduction

Financial crises appear correlated across markets or countries. As a consequence, there has been a growing interest in “*contagion*”, broadly defined as the transmission of shocks (or crises) across markets (or countries). Crises may be transmitted in two qualitatively different ways: either through stable cross-country linkages (or channels) or through sudden changes (or shifts) in these linkages of varying persistence.

From a policy perspective, it is important to discriminate between these two alternative transmission mechanisms. Short-term “insulation” policies through public sector intervention in the economy may be desirable and effective in the presence of temporary shifts in the transmission mechanism, but may not be the best (nor even a viable) policy response in presence of stable but strong linkages or a permanent change in the transmission mechanism. For instance, the temporary effects of a crisis in a neighboring country on the local foreign exchange market might call for a currency defense by means of interest rates or official reserves. But if the foreign exchange market reaction reflects more permanent features, such as strong trade and financial linkages or a long-term shift in the transmission mechanism of shocks, it is unlikely that such a defense would be advisable.

In this paper, we narrow the scope of a contagion definition well known in the literature and focus on measurement problems, with a view to distinguish between changes in cross-markets linkages during a crisis on the one hand, and strong but stable cross-markets linkages and permanent shifts in these linkages on the other hand. Specifically, we define contagion as a “*temporary* shift in the linkages across markets following a shock in one or more markets”. We then show that a time-varying coefficient model may be used to measure contagion so defined without knowing the *timing* of the crisis and in the joint presence of *heteroskedasticity* and *omitted variables*. This is achieved by (i) modelling cross-market linkages empirically as changing randomly all the time, (ii) estimating the time profile of these links with a numerical Bayesian procedure that corrects for possible omitted variable bias, and (iii) by looking at quantitatively sizable and economically plausible temporary shifts in the estimated links. Finally, the performance of the proposed measurement method is assessed by means of both simulated and actual data.

The contagion definition we adopt is that proposed by Forbes and Rigobon (2001, 2002), (henceforth, FR), and also used by King and Wadhvani (1990) among others. FR define contagion as a “significant change in cross-country linkages following a crisis in one or more countries” and call

this “shift-contagion”. As known, a strong association between two markets, both before and after a crisis in one market, is not an instance of “shift-contagion” but of “interdependence” according to this definition. We narrow the scope of this intuitive definition, by requiring that the shift in the linkages is *temporary*, to distinguish “contagion” from a permanent (or at least very persistent) shift in the transmission channels, which are usually called a “structural breaks” in the econometric literature.

Measuring contagion poses a host of statistical problems and defining it as clearly as possible is only a first step in trying to discriminate between different channels of transmission of crises across countries. In theory, one would like to use a two-steps approach to measure contagion (Favero and Giavazzi, 2002): first, by identifying the channels of transmission by estimating a model of interdependence; second, by checking whether the strength of the transmission channel has changed significantly following a crisis. However, in practice, there is a trade off between the efficiency costs of identifying all channels with large models (we shall call this full information methods) and the potential bias deriving from omitting relevant variables, observable or latent, that may distort the analysis in smaller set ups (we shall call this limited information methods).

There are several existing approaches to measure contagion.¹ These include methods based on simple rolling correlations, OLS regressions, regressions with dummy variables, and also principal component analysis. Typically, once assumed that a particular market or country is the source of the crisis, the empirical model is estimated before and after the crisis period or including dummy variables for the crisis period. Then, the statistical significance of the dummy variables or the statistical significance of the estimated differences in the coefficients before and after the crisis, is checked. Thus, all these methods assume that both the *source* and the precise *timing* of the crisis is known. This is a drawback, especially for the analysis of those crises that are difficult to date clearly, as in the case of Argentina and Turkey in 2001 and Brazil in 2002.

There are also other problems in measuring contagion of a more statistical nature.² In a limited information setting, cross-market correlations may shift even without a shift in the underlying linkages when volatility increases in the crisis country, and this (upward) bias can be corrected only if we do not have simultaneity and/or omitted variables. OLS-based and principal component methods can be safely applied in the absence of simultaneity and omitted variable problems, with

¹For a survey of the recent literature, see Pericoli and Sbracia (2002).

²For a formal discussion, see, for instance, Rigobon (2001).

the advantage that they provide also evidence on the specific channels through which shocks or crises are transmitted across markets (e.g., trade, finance, investors preference and technology, etc.). However, in the joint presence of heteroskedasticity and either omitted variables or simultaneity, these methods are biased, and inconsistent in the presence of simultaneity. Moreover, under these circumstances, there are no simple corrections that can be implemented, as extensively documented by Rigobon (2001). Finally, in a full information setting, some of the relevant variables may not be available if they are unobservable (e.g., global risk aversion).

The contagion measure we propose has a number of advantages over existing methods: the measurement method we propose works in the joint presence of heteroskedasticity and omitted variables, as the limited information approach proposed by FR but unlike OLS and principal components methods; unlike any of the existing methods discussed, there is no need to define a “crisis period” outside the empirical model, as coefficients are allowed to change continually. More generally, the framework we propose allows for analysis of both interdependence and contagion, as full information specifications are more easily estimated without running into overfitting problems using Bayesian procedures. It may distinguish between temporary shifts and structural breaks, as well as positive from negative contagion.

We apply the proposed framework to both artificial and actual data. With artificial data, we find that it detects false positives even in the most adverse experimental conditions. With actual data, we find that it replicates the results obtained in a fuller information setting when applied in a limited information setting correcting for omitted variable bias. The proposed framework can be easily implemented in a multivariate context, but may involve significant computing costs if the number of markets or countries considered is large.

The paper is organized as follows. Section two presents the econometric framework proposed and discusses its main features and properties. Given its importance in this context, the problem of omitted variable bias is dealt with separately in Section three in which we present and discuss a correction for omitted variable bias. Section four analyzes the performance of the overall framework proposed by using both artificial and actual data. Section five concludes. Some technical details of the estimation procedure used are provided in appendix.

2 Modeling Contagion and Interdependence

In this section we present a general econometric model that may be used to measure both contagion and interdependence and discuss its specification and estimation.

The transmission of shocks or crises across markets or countries, either through stable channels and linkages or through shifts or changes in these links, may be modelled by means of a standard vector-autoregression (VAR) with time-varying coefficients:

$$A_t(L)Y_t = B_t(L)W_t + D_t + U_t, \quad (1)$$

where $Y_t = [y_t^1, \dots, y_t^n]'$ is a $n \times 1$ vector of asset prices or quantities, $W_t = [w_t^1, \dots, w_t^m]'$ is a $m \times 1$ vector of controls and sources of shocks, $A_t(L)$ and $B_t(L)$ are respectively $(n \times n)$ and $(n \times m)$ time-varying polynomial matrices in the lag operator L with lag length p and q respectively, and D_t is a $n \times 1$ vector of constants. $U_t = [u_t^1, \dots, u_t^n]'$ is a $(n \times 1)$ vector of country or market specific shocks with variance-covariance matrix Σ . Thus, in principle, this specification allows for both interdependence and contagion: a stable association between two markets before and after a crisis may be traced in the usual manner through impulse response analysis, while contagion can be detected by a temporary shift in the model parameters.

This approach to the measurement of contagion has other advantages. First, as coefficients are allowed to change randomly all the time, we do not require knowledge of the precise timing of the crisis. Second, as in the case of OLS-based methods, it may provide evidence on the specific channels of transmission of shocks across markets and is not biased by shifts in volatility alone. Third, as we shall discuss in section three, unlike OLS-based methods, the approach may be adjusted to take possible omitted factors into consideration. Fourth, potential simultaneity problems may be resolved either by focusing at the variance-covariance matrix of the reduced form residuals (Σ) rather than on the estimated coefficients, or by modeling Σ as in the structural VAR literature.³

In practice, one estimates parameter values for all time observations and then look at the time profile of this series for sizable temporary shifts. As estimation is Bayesian, there is a lesser need to test the statistical significance of any economically significant shift identified. This is because the posterior distribution of the parameter of interest already summarizes the uncertainty around the point estimate, as opposed to one draw from such a distribution under a classical approach.

³See Ciccarelli and Rebucci (2001) and Coogley and Sargent (2002) for examples of structural time-varying coefficient models.

The analogous of a classical test for parameter stability, however, could be easily implemented.

2.1 Specification

Collect Y_t and W_t with all their lags and the constant term in X_t and all parameters in β_t . Then the model may be rewritten as:

$$Y_t = X_t\beta_t + \varepsilon_t, \quad (2)$$

where X_t and β_t have dimension $n \times k$ and $k \times 1$ respectively, with $k = np + mq + 1$, while Y_t and ε_t are $n \times 1$ vector stochastic processes.

To fit (2) to the data, following Canova (1993), we assume, for all t :

- (i) $\varepsilon_t \mid X_t \sim iid$ with $E[\varepsilon_t \mid X_t] = 0$ and $E[\varepsilon_t\varepsilon_t' \mid X_t] = \Sigma$;
- (ii) $\beta_t = G\beta_{t-1} + F\beta_0 + H\zeta_t$ with $\zeta_t \sim iid N(0, \Phi)$;
- (iii) X_t , ε_t and ζ_t are conditionally independent.

In addition, innovating upon Canova (1993), we assume that:

- (iv) $\varepsilon_t \mid X_t \sim iid t_\nu(0, \Omega)$, with $\Omega = \frac{\nu-2}{\nu}\Sigma$ and $\nu > 2$ (so that $E[\varepsilon_t \mid X_t] = 0$ and $E[\varepsilon_t\varepsilon_t' \mid X_t] = \Sigma$).

Here, $E[\cdot]$ is the expectation operator, “ $\sim iid$ ” means identically and independently distributed, and $N(0, \Phi)$ denotes a multivariate normal distribution with zero mean and variance-covariance matrix Φ , $t_\nu(0, \Omega)$ a centered multivariate t-student distribution with ν degrees of freedom— $\nu \in (0, \infty)$ —and (symmetric and positive definite) scale matrix Ω , while G , F , and H are known matrices of conforming dimension.

The first assumption is standard for stationary time series. The second assumption specifies the (stochastic) law of motion of the parameter vector as a general class of VAR process—including VAR processes with discrete regime shifts *a-la* Hamilton, as for instance used by Sims (1999), or the kind of process specified by Cogley and Sargent (2002). The third assumption is also standard and helps keeping the model as simple as possible, but could be relaxed in principle. The fourth hypothesis generalizes the more common $iid N(0, \Sigma)$ assumption for the vector of error terms and takes the likely presence of outliers in high frequency data into account.

In the latter regard, note first that assuming $\varepsilon_t \mid X_t \sim t_\nu(0, \Omega)$ is equivalent to assume $\varepsilon_t = \sqrt{h_t}u_t$ with $u_t \mid X_t \sim N(0, \Omega)$ and $h_t \mid X_t \sim Inv-\chi^2(\nu, 1)$, where $Inv-\chi^2(\nu, 1)$ denotes an inverted

chi-squared distribution with ν degrees of freedom and unit scale. Thus, if $\varepsilon_t | X_t \sim t_\nu(0, \Omega)$, then $\varepsilon_t | X_t, h_t \sim N(0, h_t \Sigma)$. Second, note that the t-student assigns higher probability mass on the tails of the distribution of the vector of error terms than the normal—i.e., higher probability on extreme values or outliers—and the extent to which $\varepsilon_t | X_s$ departs from normality depends on the number of degrees of freedom, ν . In fact, $\varepsilon_t | X_s$ converges in distribution to $N(0, \Sigma)$ as ν approaches infinity as in the limit $E[h_t | X_t]$ tends to one and its variance, $V[h_t | X_t]$, tends to zero.

Substitute assumption (ii) in (2) and take the conditional expectation with respect to the distribution of X_t under (i)-(iii), then we have:

$$Y_t = X_t \tilde{\beta}_{t-1} + \tilde{\varepsilon}_t,$$

where

$$\tilde{\beta}_{t-1} = G\beta_{t-1} + F\beta_0 \quad \text{and} \quad \tilde{\varepsilon}_t = X_t H \zeta_t + \varepsilon_t,$$

with

$$E[Y_t | X_t] = X_t \tilde{\beta}_{t-1} \quad \text{and} \quad V[Y_t | X_t] = \Sigma + X_t H \Phi H' X_t'.$$

Thus, under assumptions (i)-(iii), Y_t is a conditionally heteroskedastic process, with non-linear conditional mean and variance (in the vector of variables X_t). Further, under assumption (iv), Y_t is a non-normal process (i.e., with fat tails). Hence, despite its simplicity, this specification captures many typical features of high frequency financial data.⁴

2.2 Bayesian Estimation

Although simple versions of (2) under assumptions (i)-(iv) could also be estimated in a classical fashion (e.g., by using the Kalman filter, rolling regressions, or other recursive procedures), a Bayesian approach allows to estimate more general specifications for a significant number of equations.⁵ As we shall discuss below, a Bayesian approach also allows to correct for the presence of omitted variables in a quite simple manner, while a classical procedure would not allow to do so.

Bayesian estimation is simple in principle, though may be computationally demanding. Prior distributions are assigned to the hyperparameters of the model (in our case, Σ , $\tilde{\beta}_0$, Φ , and ν), and

⁴For more details on our model's ability to fit financial, high frequency time series, see Canova (1993).

⁵For specification and estimation of a time varying SUR model, see Chib and Greenberg (1995). For extension of this model to a panel VAR framework, see Canova and Ciccarelli (2000).

are combined with the information contained in the data (in the form of a likelihood function), together with a set of initial conditions, to obtain a joint posterior distribution of the parameters of interest *via* the Bayes rule. Marginal posterior distributions are then obtained by integrating out other parameters from the joint posterior distribution.

In many applications analytical integration of the joint posterior distribution may be difficult or even impossible to implement. This problem, however, can often be solved by using numerical integration methods based on Markov Chain Monte Carlo (MCMC) simulation methods (e.g., Geweke, 2000). In this paper, we use the Gibbs sampler, which is a recursive simulation method requiring only knowledge of the conditional posterior distribution of the parameters of interest on MCMC methods (e.g., Gelfand et al., 1990).

In the rest of this subsection, we describe the specific prior assumptions suggested, discuss the posterior distributions of the parameters of interest, and show how the estimation procedure may be corrected for omitted variable bias. The derivation of the posterior distributions is reported in appendix.

2.2.1 Priors

By assuming prior independence, as customarily done, the joint prior distribution of the model parameters can be expressed as the product of the marginal priors:

$$p(\Sigma, \beta_o, \Phi, \nu) = p(\Sigma) p(\beta_o) p(\Phi) p(\nu),$$

where ‘ p ’ denotes a probability density function. On these marginal priors, as also fairly standard, we assume:

$$\begin{aligned} p(\Sigma^{-1}) &= W(\varsigma, S) \\ p(\beta_o) &= N(\beta_o^*, \Theta) \\ p(\Phi^{-1}) &= W(q, Q) \\ p(\nu) &= \text{Uniform}(2, r), \end{aligned} \tag{3}$$

where $W(\varsigma, S)$ ($W(q, Q)$) denotes a Wishart distributions with degrees of freedom ς (q) and symmetric, positive definite scale matrix S (Q). The hyperparameters of these distributions ($\varsigma, q, \beta_o^*, \text{vec}(S), \text{vec}(\Theta), \text{vec}(Q)$, and r , with $\text{vec}(\cdot)$ denoting the column-wise vectorization of a matrix) are also assumed to be known.

Denote $Y^T = (Y_1, \dots, Y_T)$ the sample data and $\psi = (\{\beta_t\}_t, \{h_t\}_t, \Sigma, \beta_o, \Phi, \nu)$ the set of parameters of interest. Given prior independence and assumption (iii) above, the joint posterior density is:

$$\begin{aligned}
p(\psi | Y^T) &\propto |h_t \Sigma|^{-T/2} \exp \left\{ -\frac{1}{2} \sum_{t=1}^T (Y_t - X_t \beta_t)' (h_t \Sigma)^{-1} (Y_t - X_t \beta_t) \right\} \\
&\times |H \Phi H'|^{-T/2} \exp \left\{ -\frac{1}{2} \sum_{t=1}^T (\beta_t - \tilde{\beta}_{t-1})' (H \Phi H')^{-1} (\beta_t - \tilde{\beta}_{t-1}) \right\} \\
&\times \prod_t h_t^{-(\nu/2+1)} \exp \left\{ -\frac{1}{2} \sum_t \frac{\nu}{h_t} \right\} \\
&\times |\Theta|^{-1/2} \exp \left\{ -\frac{1}{2} (\beta_o - \beta_o^*)' \Theta^{-1} (\beta_o - \beta_o^*) \right\} \\
&\times |\Sigma|^{-\frac{1}{2}(\varsigma-n-1)} \exp \left\{ -\frac{1}{2} \text{tr}(S) \Sigma^{-1} \right\} \\
&\times |\Phi|^{-\frac{1}{2}(q-k-1)} \exp \left\{ -\frac{1}{2} \text{tr}(Q) \Phi^{-1} \right\} \\
&\times \frac{1}{r-2}
\end{aligned} \tag{4}$$

where the first line corresponds to the likelihood function, while the others represent the prior information described above, with $\tilde{\beta}_{t-1} = G\beta_{t-1} + F\beta_0$ as before.

2.2.2 Posteriors

As known, to implement the Gibbs sampler, we need to derive analytically *conditional* posterior distributions of the parameters of interest. Given the *conditional* posterior distributions of the parameters of interest, the Gibbs sampler produces an approximation to the *joint* posterior density. Convergence of the Gibbs sampler to the true invariant distribution in our case is subject to standard, mild conditions since model (2) is a time-varying SUR with serially correlated errors (e.g., Geweke, 2000). *Marginal* posterior densities are then obtained by integrating out of these joint posterior numerically within the Gibbs sampler. Moreover, inference on any continuous function of the parameters of interest, $\mathcal{G}(\psi)$, can be constructed using the output of the Gibbs sampler and the ergodic theorem.

For example

$$E(\mathcal{G}(\psi)) = \int \mathcal{G}(\psi) p(\psi|Y) d\psi$$

can be approximated using

$$\frac{1}{\bar{L}} \left[\sum_{\ell=L+1}^{L+\bar{L}} \mathcal{G}(\psi^\ell)^{-1} \right]^{-1}$$

where ψ^ℓ is the ℓ -th draws of vector ψ , $(L + \bar{L})$ is the total number of iterations in the Gibbs sampler, and L is the number of discarded iterations.

The conditional posterior distributions needed to implement the Gibbs sampler in our model are derived in appendix. Here we focus only on the interpretation of the marginal posterior means of β_t and h_t , the shift factor in variance of the error term ε_t . In particular, defining $\psi_{-\beta} \equiv (\{h_t\}_t, \Sigma, \beta_o, \Phi, \nu)$ and $\psi_{-h} \equiv (\{\beta_t\}_t, \Sigma, \beta_o, \Phi, \nu)$, in appendix we show that:

$$\beta_t \mid X_t, Y_t, \psi_{-\beta} \sim N(\hat{\beta}_t, \hat{V}_t), \quad (5)$$

with

$$\hat{\beta}_t = \hat{\beta}_{t-1} + \hat{V}_{t-1} X_t (h_t \Sigma + X_t \hat{V}_{t-1} X_t')^{-1} (Y_t - X_t \hat{\beta}_{t-1}) \quad (6)$$

$$\hat{V}_t = \hat{V}_{t-1} - \hat{V}_{t-1} X_t' (h_t \Sigma + X_t \hat{V}_{t-1} X_t')^{-1} X_t \hat{V}_{t-1}. \quad (7)$$

while

$$h_t \mid X_t, Y_t, \psi_{-h} \sim \text{Inv-}\chi^2(\nu_t, s_t^2) \quad (8)$$

with

$$\nu_t s_t^2 = \nu_{t-1} s_{t-1}^2 + (Y_t - X_t' \beta_t)' \Sigma^{-1} (Y_t - X_t' \beta_t) \quad (9)$$

$$\nu_t = \nu_{t-1} + 1. \quad (10)$$

Consider the expression for the posterior mean of the parameter vector, $\hat{\beta}_t$, equation (6). This can be written as:

$$\hat{\beta}_t = \left[X_t' (h_t \Sigma)^{-1} X_t + \hat{V}_{t-1}^{-1} \right]^{-1} \left[X_t' (h_t \Sigma)^{-1} Y_t + \hat{V}_{t-1}^{-1} \hat{\beta}_{t-1} \right]. \quad (11)$$

This, in turn, shows that, for each t , $\hat{\beta}_t$ is centered on the OLS estimator, and is identical to the OLS estimator (and thus also to the MLE estimator) if we assume that the prior distribution is non-informative—i.e., if its prior variance is set arbitrarily large or its precision arbitrarily small ($\Phi^{-1} = 0$). The posterior mean of the parameter vector, $\hat{\beta}_t$, in (5) is as unbiased as an OLS estimate, but is more efficient if the prior information is not diffuse (i.e., it entails more than complete ‘ignorance’).

To see this, note first that (6) may be written as (11). In fact, as

$$a^{-1} - (a + b)^{-1} ba^{-1} = (a + b)^{-1},$$

the following holds:

$$\begin{aligned} & \left(h_t \Sigma + X_t \hat{V}_{t-1} X_t' \right)^{-1} \\ &= (h_t \Sigma)^{-1} - \left[\left(h_t \Sigma + X_t \hat{V}_{t-1} X_t' \right)^{-1} \left(X_t \hat{V}_{t-1} X_t' (h_t \Sigma)^{-1} \right) \right]. \end{aligned}$$

Now, substituting this in (6), we have that

$$\begin{aligned} \hat{\beta}_t &= \hat{\beta}_{t-1} + \\ &+ \hat{V}_{t-1} X_t' \left[(h_t \Sigma)^{-1} - \left(h_t \Sigma + X_t \hat{V}_{t-1} X_t' \right)^{-1} X_t \hat{V}_{t-1} X_t' (h_t \Sigma)^{-1} \right] Y_t + \\ &- \hat{V}_{t-1} X_t' \left(h_t \Sigma + X_t \hat{V}_{t-1} X_t' \right)^{-1} X_t \hat{\beta}_{t-1} \\ &= \left[\hat{V}_{t-1} - \hat{V}_{t-1} X_t' \left(h_t \Sigma + X_t \hat{V}_{t-1} X_t' \right)^{-1} X_t \hat{V}_{t-1} \right] \times \\ &\left[X_t' (h_t \Sigma)^{-1} Y_t + \hat{V}_{t-1}^{-1} \hat{\beta}_{t-1} \right]. \end{aligned}$$

But since

$$(A + BCB')^{-1} = A^{-1} - A^{-1}B \left(B'A^{-1}B + C^{-1} \right)^{-1} B'A,$$

we also have that

$$\hat{\beta}_t = \left[X_t' (h_t \Sigma)^{-1} X_t + \hat{V}_{t-1}^{-1} \right]^{-1} \left[X_t' (h_t \Sigma)^{-1} Y_t + \hat{V}_{t-1}^{-1} \hat{\beta}_{t-1} \right].$$

Now remember that

$$\begin{aligned} \hat{V}_{t-1}^{-1} &= (V_{t-1}^* + H\Phi H')^{-1} \\ &= V_{t-1}^{*-1} - V_{t-1}^{*-1} H \left(H' V_{t-1}^* H + \Phi^{-1} \right)^{-1} H' V_{t-1}^{*-1}. \end{aligned}$$

Note then that $\hat{V}_{t-1}^{-1} = 0$ whenever $\Phi^{-1} = 0$, provided H is non singular, and hence we also have that,

$$\hat{\beta}_t = [X_t X_t']^{-1} [X_t Y_t].$$

if $\Phi^{-1} = 0$.⁶

⁶Note that our posterior estimates of the model parameters at time t depend on the information of the whole sample period. In a rolling OLS estimate, instead, only the information up to period t would be used.

Consider then the expression for the posterior distribution of h_t in (8). The conditional posterior distribution of h_t also has an interesting interpretation, which helps to appreciate the role of the t-distribution in the model. As we can see from (9), the expression for $\nu_t s_t^2$, which apart from a multiplicative factor provides the posterior mean of h_t , evolves as a random walk without drift. Therefore, the assumed prior structure generates a *posterior* conditional heteroskedasticity effect of the type assumed *a priori* by Cogley and Sargent (2002). Thus, this effect allows for a permanent shifts in the innovation variance, even in a specification which does not assume it *a priori*.

As an OLS estimate, however, this estimation procedure is not robust to the possible presence of omitted variables, even though a correction for omitted variables bias can be easily implemented in our model by following Leamer (1978, Chapter 9).

3 Correcting for Omitted Variable Bias

It is well known that omitting a relevant variable in the estimation of a linear model biases the estimation results and may produce false inference, even if the omitted variables are orthogonal to the variables included in the analysis. This is because of the lack of association between the omitted and the included variables produces unbiased estimates of the coefficients, but it is not sufficient to yield an unbiased estimator of their variance.

Consider a non-stochastic linear regression function:

$$Y_t = X_t\beta + Z_t\gamma \tag{12}$$

where X and Y are $n \times 1$ and $n \times k$ matrices, respectively, Z is $n \times p$ matrix (with $p < k$) and could be unobservable, while β and γ are parameter vectors. Assume, for instance, that

$$(Z | X) = XR + \eta \tag{13}$$

where η is a vector of random variables independent of X , and R is known. Thus, the true model is

$$Y = X\beta + XR\gamma + \eta\gamma \tag{14}$$

If instead we estimate the model

$$Y = X\beta + \varepsilon, \tag{15}$$

then $\hat{\beta}$ will be biased unless $R = 0$ (omitted variables are uncorrelated with the included variables) or $\gamma = 0$ (omitted variables have no effect on Y).

However, inferences about β may be made based on Y and X alone in a Bayesian estimation framework, provided we have a (probabilistic) view about Z . To see this, assume that the true model is as in (12)-(13). The model

$$Y = X\beta + X\beta^c + \xi, \quad (16)$$

where $\beta^c = R\gamma$ and $\xi = \eta\gamma$, approximates (14) by admitting the possibility of omitted variables. The fundamental difference between (15) and (16) is that the latter includes a statement about the quality of the experiment (a prior on β^c), while the former does not. In the literature, the parameter vector β_t^c is called the *contamination vector* (or the *experimental bias*) because it summarizes the bias in the information about β due to omitted variables. The model in (15) is misspecified because it sets the contamination vector to zero. Evidently, (16) could not be estimated in a classical way because of the perfect collinearity among the regressors included, but its analysis is feasible in a Bayesian context by choosing an appropriate prior to identify β_t from β_t^c .

More specifically, following Leamer (1978), assume data normality and let the prior be normal with mean and variance, respectively,

$$E \begin{pmatrix} \beta \\ \beta^c \end{pmatrix} = \begin{pmatrix} \beta^* \\ 0 \end{pmatrix}, \quad (17)$$

$$V \begin{pmatrix} \beta \\ \beta^c \end{pmatrix} = \begin{pmatrix} N^* & 0 \\ 0 & B \end{pmatrix}^{-1}, \quad (18)$$

where N^* and B are positive semi-definite matrices. Leamer (1978, p. 295) shows that the posterior mean and variance are given by, respectively,

$$E \left(\begin{bmatrix} \beta \\ \beta^c \end{bmatrix} \mid Y \right) = \begin{pmatrix} N^* + N & N \\ N & B + N \end{pmatrix}^{-1} \begin{pmatrix} N^*\beta^* + N\hat{\beta}_{ols} \\ N\hat{\beta}_{ols} \end{pmatrix} \quad (19)$$

$$V \left(\begin{bmatrix} \beta \\ \beta^c \end{bmatrix} \mid Y \right) = \begin{pmatrix} N^* + N & N \\ N & B + N \end{pmatrix}^{-1} \quad (20)$$

with $\hat{\beta}_{ols} = (X'X)^{-1}X'Y$ and $N = X'X$. By the algebra rules of partition matrices we also have:

$$\begin{aligned} & \begin{pmatrix} N^* + N & N \\ N & B + N \end{pmatrix}^{-1} \\ &= \begin{pmatrix} D^{-1} & -D^{-1}N(B+N)^{-1} \\ -E^{-1}N(N^*+N)^{-1} & E^{-1} \end{pmatrix}^{-1} \end{aligned}$$

where

$$\begin{aligned} D &= \left(N^* + N - N(B + N)^{-1} N \right) \\ E &= \left(B + N - N(N^* + N)^{-1} N \right). \end{aligned}$$

Hence,

$$\begin{aligned} & E \left(\begin{bmatrix} \beta \\ \beta^c \end{bmatrix} \mid Y \right) \\ &= \begin{pmatrix} D^{-1} \left\{ N^* \beta^* + \left[N - N(B + N)^{-1} N \right] \hat{\beta}_{ols} \right\} \\ E^{-1} N (N^* + N)^{-1} N^* (\hat{\beta}_{ols} - \beta^*) \end{pmatrix}. \end{aligned} \tag{21}$$

The posterior mean of β in (21), as usual, is a weighted average of the prior mean (β^*) and the sample OLS estimate ($\hat{\beta}_{ols}$). However, the weight of the latter is $(N - N(B + N)^{-1} N)$ rather than N as it usually happens in the absence of such a correction. Thus, the corrected estimate weights the OLS estimate less than in a model without correction. Also note that the “discount factor”, $N(B + N)^{-1} N$, depends on B (the prior precision of β^c). Hence, as B grows, the posterior mean converges to its value in a model without correction.

The posterior mean of β^c in (21) is a weighted average of *zero* and $(\hat{\beta}_{ols} - \beta^*)$, the difference between the OLS estimate and the prior mean. Hence, the posterior distribution of β^c is centered away from zero. This corrects for the excess of skewness toward β^* in the posterior distribution of β , compared to the case in which there is no correction in the model. In fact, if the posterior distribution of β^c were centered on zero and the weight of $\hat{\beta}_{ols}$ in (21) was discounted by $N(B + N)^{-1} N$, we would overweight β^* . To correct for this distortion induced by the correction, the posterior mean of β^c must be different from zero and depend on the excess of $\hat{\beta}_{ols}$ over β^* .⁷

Leamer’s correction for omitted variable bias was designed for a standard linear regression model in which the omitted variable depends on the variable included in the regression. However, it can be easily adapted to our time-varying, non-normal model, or to cases in which the omitted variable is a common factor as often assumed in the contagion literature (see, for instance, Rigobon, 2001). To adapt the correction to a time-varying model in which the omitted variable is a common factor, the prior of the parameter vector can be expressed as:

$$\delta_t = \tilde{\delta}_{t-1} + \kappa_t$$

⁷For more details, see Leamer (1978, page 297).

where

$$\begin{aligned}\tilde{\delta}_{t-1} &= \begin{pmatrix} \tilde{\beta}_{t-1} = G_1\beta_{t-1} + F_1\beta_0 \\ \tilde{\beta}_{t-1}^c = G^c\beta_{t-1} + F^c\beta_0 \end{pmatrix}, \\ \kappa_t &= \begin{pmatrix} H_1 & 0 \\ 0 & H^c \end{pmatrix} \begin{pmatrix} \zeta_{1t} \\ \zeta_t^c \end{pmatrix} = H\zeta_t,\end{aligned}$$

with

$$V(\zeta_t) = V \begin{pmatrix} \zeta_{1t} \\ \zeta_t^c \end{pmatrix} = \begin{pmatrix} \Phi_1 & 0 \\ 0 & \Phi^c \end{pmatrix} = \Phi.$$

Thus, the model (2) becomes

$$Y_t = W_t\delta_t + \epsilon_t$$

where $W_t = [X_t \ X_t]$ and $\delta_t = [\beta_t \ \beta_t^c]$. Then, the the joint posterior distribution of the parameters is given by (4), after replacing X_t with W_t , β_t with δ_t , and ϵ_t with ϵ_t .⁸

The intuition of why Leamer’s correction works also in cases in which the omitted variable is a common factor is simple. The correction exploits the correlation between the included and the excluded variables in the true model and may be interpreted as an instrumental variable estimate that uses the included regressor as instrument for the omitted regressor. For this purpose, it does not matter whether the omitted variable is common to both the dependent and the independent variable, assuming it is not endogenous to the dependent variable. We also conjecture that, when the omitted variable is a common factor, its performance might improve with the number of variables included in the model. This suggests potential scope for combining common factor analysis with Bayesian estimation methods to improve upon its performance.

4 How Does the Proposed Measure Perform?

To assess the performance of the measurement method proposed, in this section, we run two set of experiments. The first set is based on artificial data and thus a known data generating process (DGP). Here we analyze the performance of the proposed measure in a worse-case, false-positive example. The second set is based on actual data and thus an unknown DGP. Here we revisit an actual application in which both contagion and interdependence were detected and assess the performance of our measure when it is applied omitting an important source of interdependence but, at the same time, correcting as discussed in the previous section. As we shall see, our procedure

⁸The block diagonality of the variance-covariance matrix of ζ_t is a necessary prior identification assumption, but does not need to be preserved *a posteriori*.

turns out to perform remarkably well in both cases: it identifies false positives even in the most adverse experimental conditions and replicates results obtained in a much richer empirical model specification.

4.1 Evidence Based on Artificial Data

In the first set of experiments, we consider a case in which there is both heteroskedasticity and omitted variable bias and no contagion, and ask whether our measure could instead lead us to conclude erroneously (because of the variable omission) that there is contagion. Thus, we apply our measure to a case in which (i) the true cross-market linkage remains stable over time, (ii) there is interdependence, (iii) a common shock causes volatility to increase, and (iv) the model used to measure contagion omits this common source of volatility, say because this is an unobservable variable. However (v), the estimation procedure corrects for potential omitted variable bias as discussed in the previous section.

We generate the data from the following univariate, time-invariant model, consistent with model No. 3 of Rigobon (2001):

$$\begin{aligned}
 y_t &= \beta x_t + \gamma z_t + \varepsilon_t & (22) \\
 x_t &= \delta z_t + u_t, \\
 z_t &= \rho z_{t-1} + \eta_t, \\
 t &= 1, \dots, 200.
 \end{aligned}$$

In this model, the omitted variable (z_t) is a factor common to the market or country assumed to be the source of the shock or crisis (x_t) and the country potentially affected by this crisis' spillovers (y_t). This common factor may be an observable variable, such as a shock in a third market, or unobservable, such as a shift in investors preferences as discussed by Kumar and Presaud (2001).

The model is parametrized in the most unfavorable manner (to our measurement procedure) by selecting the worst-case among those considered by Rigobon (2001 pages 30-31).⁹ Hence, the parameters and error terms of the model are drawn under the assumptions that:

- $\beta \sim N(\bar{\beta}, \sigma_\beta^2)$ and $\gamma \sim N(\bar{\gamma}, \sigma_\gamma^2)$ with $\sigma_\beta = \bar{\beta}/4$, $\sigma_\gamma = \bar{\gamma}/4$, $\sigma_{\beta,\gamma} = 0$, $\bar{\beta} = \bar{\gamma} = 1$, and $\delta = 1$;

⁹It would be simple, albeit time consuming, to consider other points in the parameter space and run a proper Monte Carlo simulation experiments as done by Rigobon (2001). For the purpose of verifying the maintained statement that the proposed measure of contagion is robust to the joint presence of heteroschedasticity and omitted variable bias, however, it suffices to consider the most unfavorable point of those considered by Rigobon.

- $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ and $u_t \sim N(0, \sigma_u^2)$ with $\sigma_\varepsilon = \sigma_u = 1$ and $\sigma_{\varepsilon, u} = 0$;
- $\eta_t \sim N(0, \sigma_{1,\eta}^2)$ for $t = 1, 100$ and $\eta_t \sim N(0, \sigma_{2,\eta}^2)$ for $t = 101, 200$, with $z_0 \sim N\left(0, \frac{\sigma_\eta^2}{1-\rho^2}\right)$, $\sigma_{1,\eta} = 1$, $\sigma_{2,\eta} = \sqrt{10}$, and $\rho = 0.5$.

We then estimate this model with our time-varying procedure, omitting z_t from the first regression above, with and without Leamer's correction.

The model estimated without correction is:

$$y_t = \beta_t x_t + \varepsilon_t. \quad (23)$$

In this case, the prior assumptions for β_t and ε_t and the required initial conditions, consistent with assumptions (i)–(iv) in section 2, are:

- $G = H = I$, $F = 0$ and $\Phi = \phi V_o^*$ with $\phi = 0.001$;
- $\varepsilon_t | h_t \sim N(0, h_t \sigma^2)$ with $\sigma^2 = \hat{\sigma}_{ols}^2$;
- $h_t \sim Inv - \chi^2(\nu, 1)$ with $\nu = 5$;
- $\beta_o \sim N(\beta_o^*, V_o^*)$ with $\beta_o^* = \rho^*$ and $V_o^* = \sigma^2 (X'X)^{-1} * 10^2$, where ρ^* is the sample conditional correlation coefficient corrected as suggested by FR.

In this case, the OLS bias is given by $\gamma \frac{\delta V(z)}{\delta^2 V(z) + V(u)}$, which is increasing in $V(z)$ and decreasing with $V(u)$. If these variances change in turmoil periods, we can expect the bias to change accordingly, thus erroneously revealing presence of contagion when in fact the cross-market linkages have not changed. Volatility may shift because either $V(z)$ or $V(u)$ change. In our example, we focus on changes of $V(z)$. Therefore, we expect that our estimate of β_t is biased, with a larger bias in the second part of the sample ($t \in [100, 200]$), following the increase in the variance of η_t , erroneously leading the analyst to detect presence of contagion.¹⁰

Baig and Goldfajn (2000) note that increased volatility in the crisis country may be seen as the source of “contagion”, and the consequent strengthening of cross-market correlations even in the absence of a shift in the underlying relations is part of the “contagion” process. In this case, cross-market correlations continue to provide useful information, even though they cannot be used to

¹⁰Note that an increase in $V(u)$ decreases the bias, thus potentially leading to erroneously detect presence of positive contagion.

disentangle a shift in the linkage from other reasons for the increased co-movement across markets following a crisis. In our view, this perspective is more appealing to portfolio managers than policy makers. From a portfolio management standpoint, what matters is the extent to which asset prices comove regardless of the reasons why they do so. From policymaking standpoint, instead, it is certainly important to be able to discriminate among different sources of fluctuations in asset prices.

The model estimated with Leamer’s correction is:

$$y_t = \beta_t x_t + \beta_t^c x_t + \varepsilon_t. \quad (24)$$

In this second case, we expect that the posterior estimate of β_t is not biased (and hence β_t does not change following the increase in $V(z)$) and we assume:

- $G = G^c = I$, $F = F^c = 0$, $H = H^c = I$, $\Phi_1 = \phi V_o^*$ and $\Phi_c = \phi V_o^c$ with $\phi = 0.001$;
- $\beta_o \sim N(\beta_o^*, V_o^*)$ and $\beta_o^c \sim N(0, V_o^c)$ with $\beta_o^* = \rho^*$, $V_o^* = \sigma^2 (X'X)^{-1} \cdot 10^2$, $V_o^c = \sigma^2 I$, and $cov(\beta_o, \beta_o^c) = 0$.

Figure 1 reports the posterior mean and a 68-percent band of the estimated posterior distribution of β_t for the model estimated without correction as in (23). Figure 2 plots the results in the case in which we estimate the model with Leamer’s correction. For each sample observation, as already noted, the mean of the posterior distribution (i.e., the central line in these plots) may be compared to a rolling OLS estimate. The band contains 68 percent of the probability mass under the estimated posterior distribution of β_t and may be compared to a one-standard deviation, classical confidence interval. Thus, when the posterior mean at time t moves outside its 68-percent band at time $t - 1$, we can assume this is a statistically significant shift, in a ‘classical’ sense.

As we can see from Figure 1, when the model is estimated without correction, the posterior mean is severely biased (on average by more than 50 percent of the true value), thus not only providing a potentially misleading assessment on the presence of contagion, but also of the extent to which these two markets co-move in all states of nature. The variability of the omitted variable also induces a marked, seemingly random time-variation in the posterior mean of β_t that makes it even harder to draw conclusions. Then, as expected, the shift in the variance of the omitted, common variable at $t = 100$ produces an upward shift in the estimated coefficient of about 20

percent. This pushes the lower band of the posterior distribution above its upper bound before the shift, possibly leading the analyst to conclude that this could be evidence of contagion.

As we can see from Figure 2, the Leamer’s correction works remarkably well: (i) it reduces the bias, which on average is now only about 5 percent of the true value; (ii) it removes the random movements in the estimated posterior mean of β_t due to the omitted movements of z and (iii), most importantly, it also eliminates the shift in the coefficient due to the shift in the bias. This suggests that our proposed measure of contagion detects false positive effectively, even under rather adverse conditions.

4.2 Evidence Based on Actual Data

In the second set of experiments, we assess how the framework proposed to measure contagion works when we don’t know the true DGP. We do so by revisiting an application of the framework proposed here by one of us to the investigation of contagion from the Argentine crisis on the Chilean foreign exchange market in 2001 (Rebucci, 2002). That study concluded that, once controlled for a comprehensive set of other factors, fundamental linkages between Chile and Argentina were not strong enough to explain the peso/dollar rate in the second part of 2001, and that the presence of contagion could not be ruled out. Here, we shall *omit* all the control variables used in that study and apply the Leamer’s correction to see whether a *corrected* “limited information” model yields the same results as a “fuller information” model.

More specifically, we use two empirical models. A “full” information model, which considers the same comprehensive set of potential explanatory factors used by Rebucci (2002), and a “limited” information model, which includes only two variables, as in the experiments with simulated data discussed before, and as one would have to do in an actual multi-country application because of lack of degree of freedoms. We then estimate both models with and without Leamer’s correction for omitted variable bias and then compare the results. This should permits to see clearly the extent to which our proposed framework, when applied in a limited information setting with correction for omitted variable bias, replicates the results of a fuller information model.

The application we consider is interesting for several reasons. First, because it is a natural experiment in which both an approximate full and a limited information model can actually be specified. Chile is relatively small, even compared to other Latin American countries; there are no evident endogeneity problems, and it is possible to consider a large set of potential explanatory

factors in a single equation model.

Second, this is a case in which other measurement approaches would be difficult to apply. The Argentine crisis unfolded slowly and was far from over by the time the sample period used ended (i.e., January 2002). It would have been hard to define the right estimation window for a “before and after crisis” approach. Even assuming a window of interest could have been established, there probably would have been too few observations for efficient estimation after the crisis, while our method can be applied in real time. For the same reasons, selecting a suitable number of dummy variables could also have been difficult.

Finally, it is also an interesting case from a policy standpoint. On the one hand, the Chilean peso depreciated sharply in 2001, and there was no consensus view on which were the main driving forces. The fall in the copper price (one of Chile’s most important exports), the relative loosening of domestic monetary policy, fundamental trade and investment linkages with Argentina, and also contagion from Argentina have all been considered by financial commentators and policy analysts.¹¹ On the other hand, the central bank of Chile intervened in the foreign exchange market in August-December 2001 for the first time since the free flotation of the peso in September 1999, motivating its decision by invoking “exceptional circumstances” consistent with its previously stated intervention policy. In addition, Rebucci (2002) did not control for possible omitted variable bias. It is therefore also interesting to see whether his finding of contagion, which lends support to the central bank’s decision to intervene, would have survived controlling for such a possibility.

The full information model is the following auto-regressive distributed lag (ADL):

$$DLe_t = \alpha_t^0 + \alpha_t^1 DLe_{t-1} + \mathbf{Z}_t' \gamma_t + \varepsilon_t,$$

where e_t denotes the nominal exchange rate vis-a-vis the US dollar, $Dx_t = x_t - x_{t-1}$, $Lx_t = \log(x_t)$, and \mathbf{Z}_t represents a comprehensive set of potential explanatory variables, as listed and explained in Table 1. These include (i) a terms of trade variable (the copper price), (ii) a set of domestic factors (i.e., a set of return differential with US comparable assets), (iii) a set of regional factors (Argentine and Brazilian country and currency risk indicators and their nominal exchange rate vis-a-vis the US dollar), and finally (iv) a set of global factors (the dollar/euro rate and a semiconductor price index).¹²

¹¹See Rebucci (2002) for more details on the context of the experiments we run.

¹²We report a correlation matrix and summary statistics for all time series used in the analysis in Tables 2 and 3, respectively.

All experiments with actual data are based on the same sample period and use daily data from June 2, 1999 to January 31, 2002. This sample includes 641 observations obtained by taking only common trading days across different markets.¹³

Although this is a fairly comprehensive list, the full information model considered may still omit relevant variables, observable or unobservable. These might include, for instance, other terms of trade variables and domestic factors (such as the oil price, the long-run equilibrium relation with copper, a Chilean corporate bond spread, and a direct measure of central bank intervention), regional factors (such as Mexico, the only other investment grade country in the region), global factors (such as US corporate bond spreads and a stock return differential with the NASDAQ), and unobservable variables such as global risk aversion or global liquidity conditions. Thus, there is plenty of scope for potential omitted variable bias.

The limited information model we consider is an ADL including only a (one-trading-day) lag of the exchange rate log-change and the contemporaneous change in the Argentine country spread:

$$DLe_t = \alpha_t^0 + \alpha_t^1 DLe_{t-1} + \gamma_t Di_t^{AR} + \varepsilon_t.$$

Thus, the second model omits all control variables included in the first model, and particularly two (observable) common factors between Chile and Argentina found to have significant explanatory power by Rebucci (2002): the Brazilian country risk indicator and nominal exchange rate (see also, for instance, the correlation matrix in Table 2). In fact, this second model is analogous to a rolling-correlation or rolling-OLS analysis, except for the lagged endogenous variable included to capture some predictability detected in the data (result not reported). We note finally that this is the same specification one would likely want to adopt in a multi-country application because of the need in that case to keep the model as parsimoniously parametrized as possible.

Defining $y_t \equiv DLe_t$ and collecting right-hand-side variables of both the full and the limited information model in x_t we have:

$$y_t = x_t' \beta_t + \varepsilon_t. \tag{25}$$

¹³As discussed by Rebucci (2002), the first difference of the level, or the log-level, of the variables are calculated with respect to the previous trading day included in the sample. By proceeding in this manner, consistency across variables at any given point in time is assured. Because of this, however, the first difference following a holiday may refer to more than one trading day. This potentially creates outliers artificially. Alternatively, observations following non-overlapping holidays would reflect different information sets across variables and time. Either way, we would introduce some noise into the data. Given that the estimation procedure used is robust to the presence of outliers, the former approach is preferable.

For both the full and limited information model, the prior assumptions for β_t and ε_t , and the required initial conditions, consistent with the hypotheses (i)-(iv) in section 2 and in (3), are:

- $G = H = I$, $F = 0$ and $\Phi = \phi V_o^*$ with $\phi = 0.001$;
- $\varepsilon_t | h_t \sim N(0, h_t \sigma^2)$ with $\sigma^2 = \hat{\sigma}_{ols}^2$;
- $h_t \sim Inv - \chi^2(\nu, 1)$ with $\nu = 5$;
- $\beta_o \sim N(\beta_o^*, V_o^*)$ with $\beta_o^* = \rho^*$ and $V_o^* = \sigma^2 (X'X)^{-1} * 10^2$.

Both models are then estimated with and without Leamer's correction. With correction, we initialize the model in three steps. First, we estimate the model

$$y_t = x_t' \beta_t + x_t' \beta_t^c + \varepsilon_t, \quad (26)$$

without time variation, specifying the prior assumptions (19) and (20), assuming $\beta^{c*} = 0$, $\beta^* = \rho^*$, $N = \sigma_{ols}^2 (X'X)^{-1}$, $N^* = \varphi_1 N$, and $B = \varphi_2 (N^* + N)$, and estimating the two hyperparameters φ_1 and φ_2 by maximizing the likelihood of the data. Second, we initialize β_o^* , β_o^{c*} , V_o^* , and V_o^{c*} in the time varying model with the time-invariant posterior mean of β and β^c and their variance-covariance matrices, as we do with artificial data. Finally, we assume $G = G^c = I$, $F = F^c = 0$, $H = H^c = I$ and $\Phi_1 = \Phi_c = \phi V_o^*$, setting ϕ to an arbitrarily small number (i.e., 0.001), as commonly done in the literature.¹⁴

Without correction, we set V_o^* and σ^2 equal to the OLS estimates of (25) assigning the correlation coefficient corrected for the presence of heteroskedasticity by FR as the prior mean of β_o , β_o^* .

In both the corrected and not corrected specification we set ν_o to 5. The Gibbs sampler then iterates 5000 times and discards the first 2500 draws to guarantee independence from initial conditions. We check for convergence by calculating the mean of the draws for 500, 1000, 1500, 2000 observations respectively and find that convergence is achieved after the first 1000 observations.

Figure 3 reports the posterior mean of γ_t and 68-percent bands of the estimated posterior distribution for each trading day in the sample, in all cases considered. To provide a benchmark for comparison and to help assess the results, Figure 3 also reports an 80-day rolling correlation

¹⁴Setting ϕ arbitrarily small implies assuming relatively little parameter time-variation, a priori. However, a proper prior assumption could also be given to ϕ to increase the efficiency of the estimates obtained.

between the log-change of the Chilean peso and the change in the Argentine country risk indicator (upper, left panel) and a plot of their *levels* (upper, right panel).

The results for the *not corrected, full information model* (lower, left panel) (the same result reported by Rebucci, 2002) show clear evidence of a temporary change in the linkage between these two countries, and thus indicate the presence of contagion according to the definition adopted. In fact, we can clearly see a temporary increase in the strength of the association between the Chilean exchange rate and the Argentine country risk indicator, and the magnitude of these changes leaves no doubt on their economic significance. The coefficient of the Argentine country risk starts to increase markedly at the beginning of July 2001 (upper, right panel), around the time the Chilean peso first jumped sharply, after the Argentine “mega-swap” failed to restore investor confidence, following some decline in the proceeding two-three months. The magnitude of this coefficient more than doubled in a few days after July 3, to reach a relative peak at about three times its end-June level on August 1, following a second downgrade of the Argentine sovereign rating in a few weeks. The coefficient reached its maximum on October 10, declining gradually thereafter, to bottom out on December 28 and revert to its per-June 2001 values in early January 2002, despite the Argentine country risk remaining at very high levels.

In the *corrected, full information model* (middle, left panel) the evidence of contagion is slightly weaker, statistically, as the lower band during the turmoil period remains below the posterior mean during the proceeding tranquil period. Nonetheless, the magnitude (and hence the economic significance) of the shift in the cross-market linkage remains impressive: the coefficient of the Argentine country risk indicator peaks during the turmoil period at about two times its value during the proceeding tranquil period, even after controlling for potential omitted variable bias. We conclude from this evidence that observed shift does not appear as the sole artefact of increased volatility in Argentina, or as the result of an estimation bias due to the omission of other factors, and thus confirms the previous finding of contagion.

The results in the case of a *not corrected, limited information model* (lower, right panel) are clearly different from those obtained in full information settings and show the large impact of the omitted variable bias on the estimated posterior distribution. As a result, had an analyst used such an approach, it would have been much more difficult to draw inference on the extent to which the Chilean foreign exchange market was affected by contagion from Argentina in 2001. Even though a strengthening of the cross-country linkage is evident also in this case, its quantitative magnitude

is greatly overstated, and it would have been quite difficult to identify when contagion actually started. Conclusions drawn from this model would have been similar to those one could have drawn based on the rolling correlation analysis reported in the upper, left panel.¹⁵

As expected, we can see that a *corrected, limited information model* (middle, right panel) performs almost as well as the corrected, full information model. There is almost no bias compared to the latter and the inference one could draw based on this evidence is the same as that one would have drawn in the fuller information settings. Thus, confirming that our contagion measure works well also when applied to actual data coming from an unknown DGP.

5 Conclusions

In this paper we have proposed to use a time-varying coefficient model estimated with a numerical Bayesian procedure to measure contagion empirically. We have also shown that this framework has a number of advantages, which are important from a policy standpoint in our view.

In particular, we have shown that it works well in the joint presence of heteroskedasticity and omitted variables bias, and hence it may be applied both in a full or limited information setting. By modeling the variance covariance matrix of the system, it could also be easily extended to a multivariate, simultaneous equations context. In addition, it does not require knowledge of the timing of the crisis and distinguishes contagion not only from interdependence but also from structural breaks. Finally it could be used to investigate positive as well as negative contagion.

Overall, the evidence we reported suggests that the proposed framework measures contagion effectively. Evidence based on a worse-case scenario generated with artificial data shows that the proposed framework effectively detects false positives in the joint presence of heteroskedasticity and omitted variable bias. Evidence based on actual data shows that the results obtained in a limited information setting correcting for potential omitted variables bias are comparable to those obtained in fuller information settings.

¹⁵Note, however, that rolling correlations are biased also in the sole presence of increased volatility in the crisis country, while this procedure, as well as OLS regressions, are not.

A Deriving Conditional Posterior Distributions

In this appendix we derive the *conditional* posterior distributions of the parameters of interest needed to implement the Gibbs sampler.

Assume a fixed number of degrees of freedom of the t -distribution of the error term, ν .¹⁶ Let $\tilde{\beta}_{t-1} = G\beta_{t-1} + F\beta_0$ and $\beta_t^\dagger = \beta_t - G\beta_{t-1}$. Recall that $\psi = (\{\beta_t\}_t, \{h_t\}_t, \Sigma, \beta_o, \Phi)$ and focus first on $\psi_{-\beta_t, h_t} = (\Sigma, \beta_o, \Phi)$.

From (4), the following three posterior distributions can be derived analytically. First,

$$\Sigma^{-1} \mid Y^T, \psi_{-\Sigma} \sim W(\hat{\varsigma}, \hat{S}), \quad (27)$$

where

$$\begin{aligned} \hat{\varsigma} &= \varsigma + T, \\ \hat{S}^{-1} &= S^{-1} + \sum_t (h_t^{-1}) (Y_t - X_t \beta_t) (Y_t - X_t \beta_t)'; \end{aligned}$$

second,

$$\Phi^{-1} \mid Y^T, \psi_{-\Phi} \sim W(\hat{q}, \hat{Q}), \quad (28)$$

where

$$\begin{aligned} \hat{q} &= q + T \\ \hat{Q}^{-1} &= Q^{-1} + \sum_t [H(\beta_t - \tilde{\beta}_{t-1})] [H(\beta_t - \tilde{\beta}_{t-1})]'; \end{aligned}$$

and third

$$\beta_o \mid Y^T, \psi_{-\beta_o} \sim N(\hat{\beta}_o, \hat{\Theta}), \quad (29)$$

where

$$\begin{aligned} \hat{\beta}_o &= \hat{\Theta} \left[\sum_t F' (H\Phi H')^{-1} \beta_t^\dagger + \Theta^{-1} \beta_o^* \right], \\ \hat{\Theta} &= \left[\sum_t F' (H\Phi H')^{-1} F + \Theta^{-1} \right]^{-1}. \end{aligned}$$

¹⁶The assumption of a fixed ν could be relaxed. In this case, the Gibbs sampler should be augmented by a step for sampling from the conditional posterior of ν . No simple method exists for this step, but a Metropolis step could be easily used instead. A complication, however, is that such models usually have multimodal posterior densities, requiring to search for all modes and jump between modes in the simulation (see Gelman et al., 1995, Chapter 12).

Since the conditioning on other parameters assumed independent is irrelevant, the first conditional posterior is obtained from the first and fifth lines of (4), the second from the second and sixth lines of (4), and the third and fourth lines of (4).

Focus now on the conditional posterior distributions of β_t and h_t , and particularly on $p(\beta_t | Y^T, \psi_{-\beta_t})$. Assume further that, *a priori*,

$$\tilde{\beta}_{t-1} \sim N(\beta_{t-1}^*, V_{t-1}^*). \quad (30)$$

Given $\beta_t = \tilde{\beta}_{t-1} + H\zeta_t$ ($\tilde{\beta}_{t-1} = G\beta_{t-1} + F\beta_0$) and (30), it follows that:

$$\beta_t | X_t \sim N(\hat{\beta}_{t-1}, \hat{V}_{t-1}) \quad (31)$$

where

$$\hat{\beta}_{t-1} = \beta_{t-1}^*, \text{ and } \hat{V}_{t-1} = V_{t-1}^* + H\Phi H'.$$

Now given the conditional normality of the data and (31), the joint conditional density of Y_t and β_t , $p(Y_t, \beta_t | X_t, h_t)$, is:

$$\begin{pmatrix} Y_t \\ \beta_t \end{pmatrix} | X_t, \psi_{-\beta_t} \sim N \left[\begin{pmatrix} X_t \hat{\beta}_{t-1} \\ \hat{\beta}_{t-1} \end{pmatrix}, \begin{pmatrix} X_t \hat{V}_{t-1} X_t' + h_t \Sigma & X_t \hat{V}_{t-1} \\ \hat{V}_{t-1} X_t' & \hat{V}_{t-1} \end{pmatrix} \right].$$

Then, by using the properties of the multivariate normal distribution, from this joint posterior distribution it is possible to compute the posterior distribution of β_t conditional on h_t , Y_t and the other parameters as:

$$\beta_t | X_t, Y_t, \psi_{-\beta_t} \sim N(\hat{\beta}_t, \hat{V}_t) \quad (32)$$

where

$$\hat{\beta}_t = \hat{\beta}_{t-1} + \hat{V}_{t-1} X_t' (h_t \Sigma + X_t \hat{V}_{t-1} X_t')^{-1} (Y_t - X_t \hat{\beta}_{t-1}) \quad (33)$$

and

$$\hat{V}_t = \hat{V}_{t-1} - \hat{V}_{t-1} X_t' (h_t \Sigma + X_t \hat{V}_{t-1} X_t')^{-1} X_t \hat{V}_{t-1}.$$

Consider now the posterior distribution of h_t , $p(h_t | X_t, Y_t, \psi_{-h_t})$. The joint density function of Y_t and h_t can be obtained as the product of the likelihood function (first line of 4) and the prior density of h_t (the third line of 4), which as noted has the form of an inverted chi-square distribution. For instance, for $t = 1$, it is

$$\begin{aligned} & |h_1 \Sigma|^{-T/2} \exp \left\{ -\frac{1}{2} (Y_1 - X_1 \beta_1)' (h_1 \Sigma)^{-1} (Y_1 - X_1 \beta_1) \right\} \\ & \times (h_1)^{-(\frac{\nu_o}{2} + 1)} \exp \left[-\frac{\nu_o}{2h_1} \right], \end{aligned} \quad (34)$$

where the second line is proportional to the density of an inverted chi-squared distribution with ν_o degrees of freedom and scale $s_o^2 = 1$. The product in (34), in turn, is proportional to

$$(h_1)^{-\left(\frac{\nu_o+1}{2}+1\right)} \exp \left[-\frac{1}{2h_1} \left(\varepsilon_1' \Sigma^{-1} \varepsilon_1 + \nu_o \right) \right],$$

which is an inverted chi-squared distribution, with $\nu_1 = \nu_o + 1$ degrees of freedom and scale s_1^2 , where

$$\nu_1 s_1^2 = \nu_o s_o^2 + \varepsilon_1' \Sigma^{-1} \varepsilon_1$$

with $\varepsilon_t = (Y_t - X_t \beta_t)$. Hence, by iterating recursively find that, for any t :

$$h_t \mid X_t, Y_t, \psi_{-h_t} \sim Inv - \chi^2 \left(\nu_t, s_t^2 \right) \quad (35)$$

with

$$\nu_t s_t^2 = \nu_{t-1} s_{t-1}^2 + \varepsilon_t' \Sigma^{-1} \varepsilon_t$$

and

$$\nu_t = \nu_{t-1} + 1.$$

The Gibbs sampler cycles through (27)–(35). To operationalize the entire procedure, one finally needs values for the hyperparameters of the model and suitable initial conditions for the parameters of (30), which in turn requires to specify the matrix V_{t-1}^* and the vector β_{t-1}^* . For instance, to derive the results in Figure 3 we set β_o^* and V_o^* equal to OLS estimates of (25), while ν_o was set arbitrarily to allow for the maximum degree of departure from normality.

References

- [1] Baig T. and I. Goldfajn (2000), “The Russian Default and the Contagion to Brazil,” IMF WP/00/160.
- [2] Canova (1993), “Modelling and Forecasting Exchange rates with a Bayesian Time-Varying Coefficient Model”, *Journal of Economic Dynamic and Control*, Vol. 17, pages 233-261.
- [3] Canova F. and M. Ciccarelli (2000), “Forecasting and Turning Point Prediction in a Bayesian Panel VAR Model,” UPF WP No AD 2000-05 (Also forthcoming in the *Journal of Econometrics*).
- [4] Chib S. and E. Greenberg (1995), “Hierarchical Analysis of SUR Models with Extensions to Correlated Serial Errors and Time-Varying Parameter Models,” *Journal of Econometrics*, 68, 409–431.
- [5] Ciccarelli M. and A. Rebucci (2001), “The Transmission Mechanism of European Monetary Policy: Is there Heterogeneity? Is It Changing Over Time?,” Bank of Spain WP No. 0115 and IMF WP No. 02/54.
- [6] Cogley T. and T.J. Sargent (2002), “Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.,” unpublished manuscript, Arizona State University.
- [7] Favero C. and F. Giavazzi (2002), “Is the International Propagation of Financial Shocks Non-Linear? Evidence from the ERM,” *Journal of International Economics*, Vol. 57, No.1, pp. 231-246.
- [8] Forbes and Rigobon (2001), “Contagion in Latin America: Definitions, Measurement, and Policy Implications” *Economia*, Vol. 1, No. 2, pp. 1-46.
- [9] Forbes and Rigobon (2002), No Contagion, Only Interdependence: Measuring Stock Market Co-Movements, *Journal of Finance*, Vol. 57, No. 5, pp 2223-2261.
- [10] Gelfand A.E., S. E. Hills, A. Racine-Poon, and A.F.M. Smith (1990), “Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling,” *Journal of the American Statistical Association*, Vol. 85, pp. 972-985.

- [11] Geweke, J. (2000), "Simulation Based Bayesian Inference for Economic Time Series", in Shuermann M. R. and M. Weeks (eds.), *Simulation Based Inference in Econometrics: Methods and Applications*, Cambridge University Press.
- [12] King, M. A. and S. Wadhvani (1990), "Transmission of Volatility between Stock Markets," *Review of Financial Studies*, Vol. 3, No. 1, pp. 5-33.
- [13] Leamer E.E. (1978), *Specification Searches. Ad Hoc Inference with Non-Experimental data*, (New York: Wiley).
- [14] Kumar M. and A. Presaud (2001), "Pure Contagion and Investors' Shifting Risk Appetite: Analytical Issues and Empirical Evidence" IMF Working paper No. 01/134.
- [15] Pericoli M. and M. Sbracia (2003), "A Primer on Financial Contagion", forthcoming in the *Journal of Economic Surveys*.
- [16] Rebucci A. (2002), "The Determinants of the Chilean Peso/US Dollar Spot Rate Under the Free Floating Regime: Is there an 'Argentine Factor?', Is there Evidence of 'Shift-Contagion'?", in *Chile: Selected Issues*, IMF Staff Country Report No. 02/163 (available at www.imf.org).
- [17] Rigobon R. (2001), "Contagion: How to measure It?" NBER WP No. 8118.
- [18] Rigobon R. (2003) "On the Measurement of the International Propagation of Shocks: Is it Stable?" forthcoming in the *Journal of International Economics*.
- [19] Sims C. (1999), "Drifts and Breaks in Monetary Policy," unpublished manuscript, Princeton University.

Table 1. The Set of Potential Explanatory Factors Considered

Acronimous	Name	Definition	Unit of Measure	Sampling	Source
DLe	Chilean spot rate	Log-change in the Chilean peso /U.S. dollar rate	Daily return in percent	Closing quote	Bloomberg
DLe	Copper price	Log-change in the London metal Exchange spot copper price	Daily return in percent	Closing quote	Bloomberg
m	Interest rate differential	Short-term interest rate differential (TAB-90 rate minus federal fund rate) (TAB-90 rate in UF)	Percentage point per year	Daily average	Bloomberg and Asociacion de Bancos
Dm	Interest rate differential change	Change in short-term interest rate differential (TAB-90 rate minus federal fund rate)	Percentage point per year	Daily average	Bloomberg
DiCHL	Chilean sovereign risk	Change in the Chilean component of the EMBI Global index	Percentage point per year	Unknown	Bloomberg
DfCHL	Chilean currency risk	Change in the differential between the implied one-year NDF interest rate and the one-year U.S. Treasury yield (constant to maturity)	Percentage point per year	Mid-yield	Bloomberg and IMF ICM Department
s	Stock market differential	Stock market daily return differential (IGPA index minus S&P500 index)	Percentage point per day	Closing quote	Bloomberg
DiAR+	Argentine sovereign risk	Change in the Argentine component of the EMBI+ index	Percentage point per year	Closing quote	Bloomberg
DfAR	Argentine currency risk	Change in the differential between the implied one-year NDF interest rate and the one-year U.S. Treasury yield (constant to maturity)	Percentage point per year	Mid-yield	Bloomberg and IMF ICM Department
DLeAR	Argentine spot rate	Log-change in the Argentine peso/U.S. dollar rate	Daily return in percent	Closing quote	Bloomberg
DiBR+	Brazilian sovereign risk	Change in the Brazilian component of the EMBI+ index	Percentage point per year	Closing quote	Bloomberg
DfBR	Brazilian currency risk	Change in the differential between the implied one-year NDF interest rate and the one-year U.S. Treasury yield (constant to maturity)	Percentage point per year	Mid-yield	Bloomberg and IMF ICM Department
DLeBR	Brazilian spot rate	Log-change in the Brazilian real/US dollar rate	Daily return in percent	Closing quote	Bloomberg
DLb	Semiconductor price	Log-change in a semiconductor spot price (DRAM module, 100 mghz bus 128 MB)	Daily return in percent	Unknown	Datastream (DRMU03S)
DLeEU	Euro spot rate	Log-change in the Euro/U.S. dollar rate	Daily return in percent	Closing quote	Bloomberg

All exchange rates are expressed in units of national currencies per U.S. dollar.

Table 2. Sample Correlation Matrix (June 2, 1999 - January 31, 2002)

	Chilean Spot Rate	Semi-conductor Price	Euro Spot Rate	Interest Rate Differential	Interest Rate Differential Change	Chilean Sovereign Risk	Chilean Currency Risk	Stock Market Differential	Copper Price	Brazilian Sovereign Risk	Brazilian Currency Risk	Brazilian Spot Rate	Argentine Sovereign Risk	Argentine Currency Risk	Argentine Spot Rate
	Dle	DLb	DLeEU	m	Dm	DiCHLg	DfCHL	s	DLc	DiBR+	DfBR	DLeBR	DiAR+	DfAR	DLeAR
Chilean spot rate	1,00														
Semiconductor price	-0,07	1,00													
Euro spot rate	0,01	0,03	1,00												
Interest rate differential	-0,06	0,15	0,01	1,00											
Interest rate differential change	-0,03	0,03	-0,01	0,11	1,00										
Chilean sovereign risk	0,01	0,00	-0,03	-0,03	-0,04	1,00									
Chilean currency risk	-0,07	0,00	0,01	-0,05	-0,07	0,06	1,00								
Stock market differential	0,09	-0,06	-0,06	0,01	-0,03	0,04	0,03	1,00							
Copper price	-0,04	0,14	-0,02	0,07	0,05	0,03	-0,02	-0,11	1,00						
Brazilian sovereign risk	0,33	-0,07	-0,07	-0,08	-0,11	0,10	0,07	0,22	-0,10	1,00					
Brazilian currency risk	0,15	-0,05	-0,01	-0,03	-0,05	0,03	0,05	0,05	0,00	0,24	1,00				
Brazilian spot rate	0,37	0,01	0,02	-0,06	-0,06	0,00	0,02	0,09	-0,08	0,41	0,18	1,00			
Argentine sovereign risk	0,19	-0,02	0,06	0,11	-0,02	0,02	-0,02	-0,01	0,01	0,36	0,14	0,19	1,00		
Argentine currency risk	0,18	-0,02	0,00	0,12	-0,03	-0,10	0,06	-0,06	0,03	0,25	0,18	0,22	0,34	1,00	
Argentine spot rate	0,05	0,03	0,02	0,22	0,00	0,02	-0,07	-0,03	0,01	0,02	0,01	0,06	0,11	-0,07	1,00

Sources: Bloomberg; Datastream; Fund database (ICM Dept.); and Fund staff calculations.

Table 3: Sample Descriptive Statistics (June 2, 1999-January 31, 2002)

	Chilean spot rate	Semi-conductor Price	Euro spot rate	Interest rate differential	Interest rate differential change	Chilean sovereign risk	Chilean currency risk	Stock market differential	Copper price	Brazilian sovereign risk	Brazilian currency risk	Brazilian spot rate	Argentine sovereign risk	Argentine currency risk	Argentine spot rate
	Dle	DLb	DLeEU	m	Dm	DiCHLg	DfCHL	s	DLc	DiBR+	DfBR	DLeBR	DiAR+	DfAR	DLeAR
Mean	0,05	-0,19	0,03	0,28	0,00	0,00	-0,01	0,05	0,02	0,00	0,00	0,05	0,06	0,21	0,10
Median	0,04	0,00	0,04	-0,04	0,00	0,00	-0,01	0,07	0,00	-0,01	-0,01	0,08	0,01	0,00	0,00
Standard Deviation	0,49	4,17	0,69	1,57	0,18	0,08	0,15	1,28	1,18	0,18	0,40	0,88	0,72	3,16	1,67
Kurtosis	2,62	27,48	3,17	3,15	19,35	15,74	4,42	1,67	6,18	1,73	5,86	4,51	45,13	65,17	266,33
Skewness	0,35	2,84	-0,56	1,88	0,02	-0,33	0,04	-0,08	0,77	0,27	0,68	-0,07	0,51	-0,60	14,26
Minimum	-1,92	-17,89	-4,47	-1,42	-1,41	-0,55	-0,72	-4,88	-4,77	-0,71	-1,96	-4,40	-7,96	-38,88	-7,84
Maximum	2,43	42,02	2,03	6,00	1,39	0,55	0,86	6,54	8,90	0,69	2,16	5,21	7,15	33,30	33,65
Number of observations	641	641	641	641	641	641	641	641	641	641	641	641	641	641	641

Sources: Bloomberg; Datastream; Fund database (ICM Dept.); and Fund staff calculations.

Figure 1. Posterior Distribution of β_t (Without Correction)

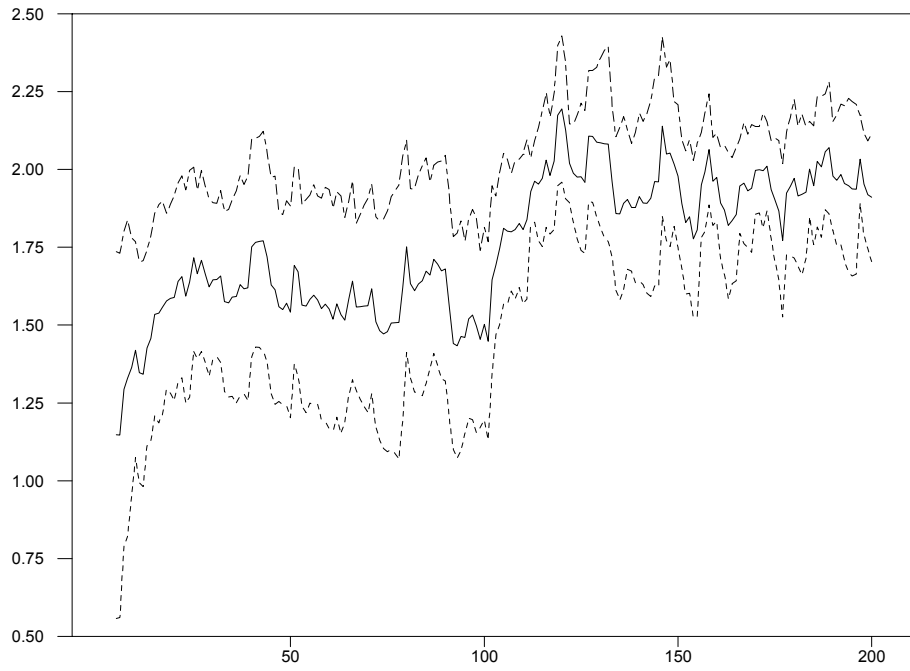


Figure 2. Posterior Distribution of β_t (With Correction)

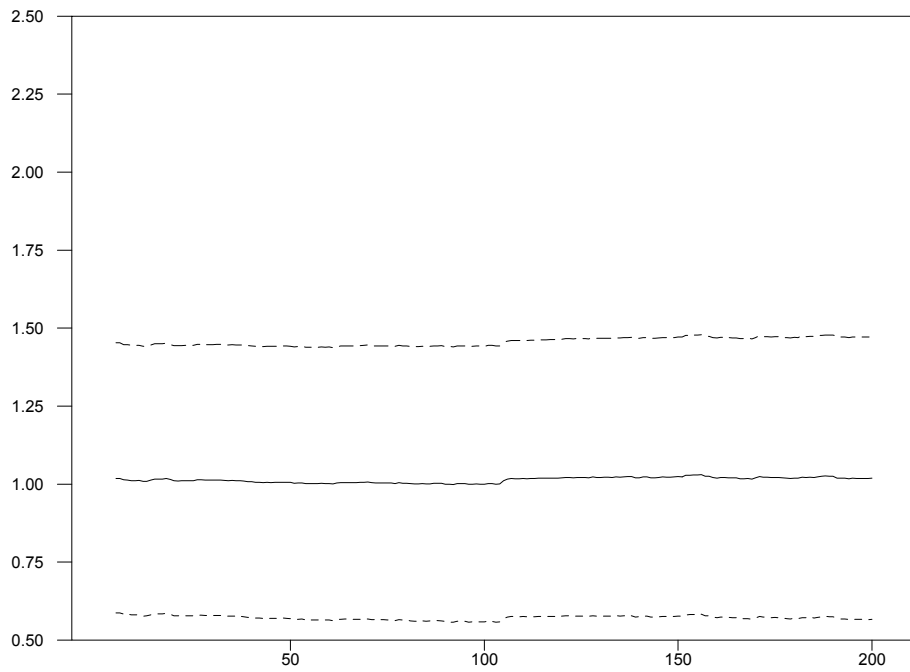
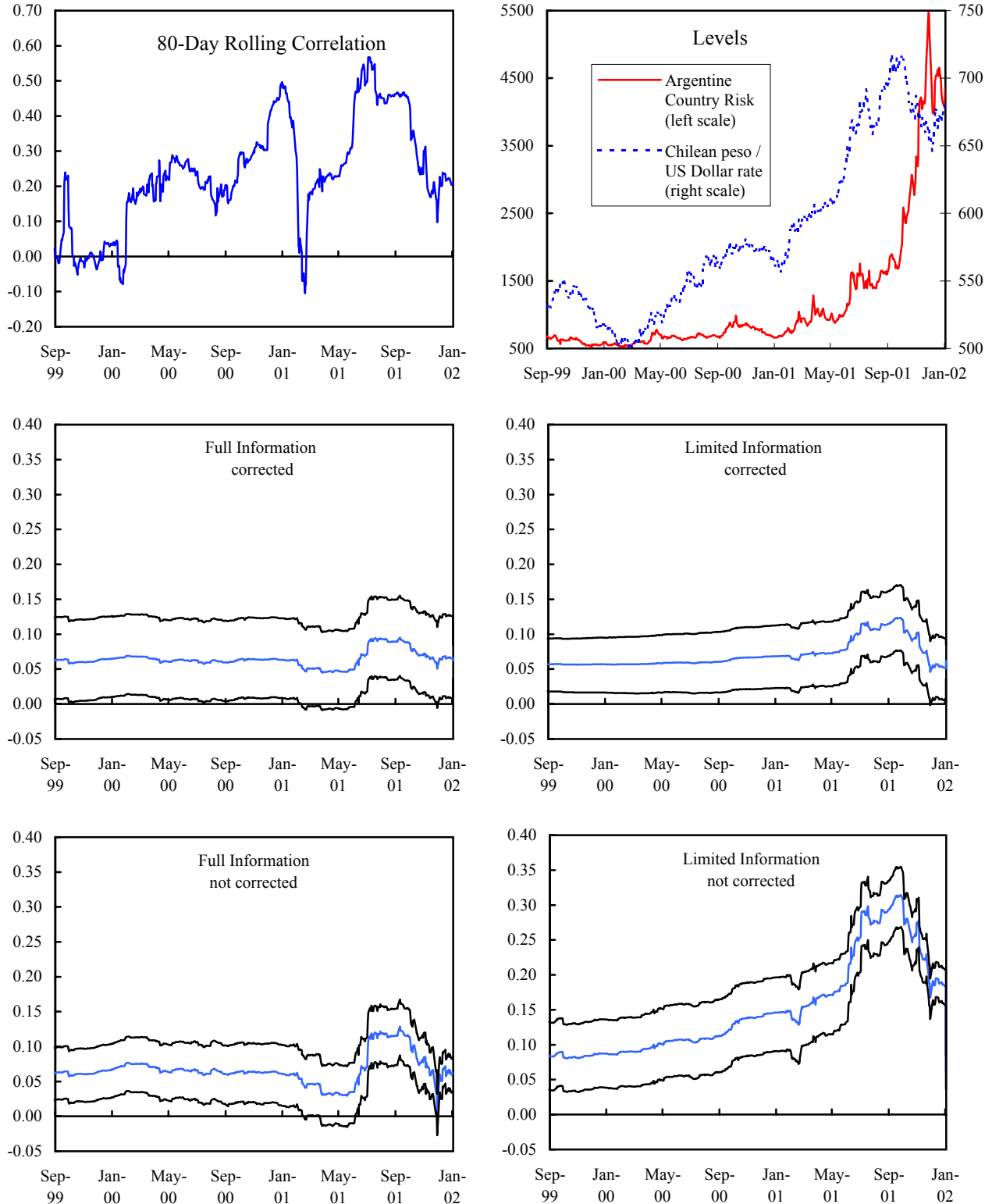


Figure 3. Alternative Measures of Contagion
From the Argentine Crisis to the Chilean Foreign Exchange Market



Source: Authors' calculations based on data described in Table 1.