The Count of Monte Carlo: Analysing Global Banking and Currency Crises, 1883-2008.

Mardi Dungey
University of Tasmania

Vance L. Martin
University of Melbourne

A.R.Tremayne
University of New South Wales and University of Liverpool

November 2012
Preliminary

Abstract

An empirical model of financial crises is specified and estimated using annual data for the period 1883 to 2008. Contagion is explicitly modelled by allowing additional linkages across countries when the number of crises reaches a threshold level. As the data are characterised as counts of the number of countries experiencing a crisis within a year, a threshold integer autoregressive moving average model (TINARMA) is specified. An EMM estimator which uses Monte Carlo simulations is proposed as estimation by maximum likelihood is infeasible. The empirical results show strong evidence of contagion during currency and banking crises, with additional evidence of a potential 10 year cycle.

Keywords: count time series; binomial thinning; efficient method of moments

JEL Classification: C150, C220, C250

1 Introduction

The most common way of modelling the transmission of financial crises and contagion is to measure the change in the covariability of asset returns across countries during a crisis period (see Dungey, Fry, González-Hermosillo and Martin (2005, 2010) for recent reviews of models of contagion). Perhaps a more intuitive way of identifying the existence of a financial crisis is simply to view a financial crisis as having a domino-effect on countries by counting the number of countries over time that move into crisis mode. This is the approach adopted in this paper where time series data on the number of countries exhibiting a financial crisis are used to model contagion. Both currency and banking crises are analysed using annual data covering the period 1883 to 2008.

As the data are characterised by low order counts of countries in crisis, with contagion occurring as an increase in correlation during the crisis period, an alternative class of models is specified to capture explicitly these characteristics. This is especially important where the model is used to generate coherent forecasts of the number of countries exhibiting a financial crisis (McCabe and Martin (2005), Jung and Tremayne (2006a)). Formally this involves specifying an integer autoregressive moving average model (INARMA) where the usual operator of the ARMA class of models is replaced by the binomial thinning operator to preserve the integer status of the dependent variable (Al-Osh and Alzaid (1988), McKenzie (1988)). To capture the presence of contagion during financial crises a threshold version of the INARMA model (TINARMA) is specified with the threshold determined by a prespecified number of countries in a financial crisis (see also Brännäs and Hellstrom (2001), Brännäs and Hall (2001), and Brännäs and Shahiduzzaman Quoreshi (2010)).

Whilst maximum likelihood estimators are available for the INAR(1) model (Al-Osh and Alzaid (1987)) as well as for higher-order INAR models (Bu, Hadri and McCabe (2008)), the inclusion of a moving average component makes maximum likelihood estimation infeasible.[1] The approach adopted here is to follow Martin, Tremayne and Jung (2012) and use a simulation based estimator. Under certain conditions this estimator is consistent and achieves the same level of efficiency as the maximum likelihood estimator. The performance of this estimator is investigated for a range of INARMA

[1] Other methods that have been proposed to estimate moving average parameters in integer models include generalized methods of moments and conditional least squares (Al-Osh and Alzaid (1988); McKenzie (1988); Brännäs and Hall (2001) and Brännäs and Shahiduzzaman Quoreshi (2010)).

models in Martin, Tremayne and Jung (2012).

The empirical results show convincing evidence of contagion during the currency and banking crises that have occurred over the period 1883 to 2008. An important feature of the empirical results is that there is strong evidence of a potential 10 year cycle in the data. The rest of the paper proceeds as follows. Section 2 provides an initial discussion of the data used to model financial crises. The TINARMA model is specified in Section 3. Section 4 provides details of the EMM estimator and the methods used to construct standard errors, which are applied in Section 5 to analysing the data discussed in Section 2. Concluding comments and suggestions for future research are presented in Section 6

2 Financial Crises: A First Look at the Data

The crisis data analysed consist of the total number, out of a group of 21 countries, that experienced some financial crisis in a particular year, over the period 1883 to 2008, a sample of size $T = 126$. The countries are: Argentina; Australia; Belgium; Brazil; Canada; Chile; Denmark; Finland; France; Germany; Greece; Italy; Japan; Netherlands; Norway; Portugal; Spain; Sweden; Switzerland; the United Kingdom; and the United States. Two types of financial crises are investigated, consisting of banking and currency crises. See Dungey, Jacobs and Lestano (2012) as well as Appendix A for a description of how the data are constructed.

The data are presented in Figure 1 which provides a time series plot of the total number of countries with banking crises on the one hand and currency crises on the other per annum over the period, as well as a combination of the two. Financial crises were most widespread in 1931 where 14 of the 21 countries in the sample experienced banking and currency crises. The next biggest crisis occurred in 1971 where 12 of the 21 countries experienced a crisis in their currency. Table 1 gives a break down of financial crises per country. Argentina experienced the most number of currency crises (20), followed by Brazil (14) and the UK (11). The US has had the most number of banking crises (10), followed by Argentina and Brazil (9 each) and Italy (8). Argentina has experienced simultaneous currency and banking crises on 4 occasions, followed by Brazil and Finland (3 each). A perusal of the panels of Figure 1 makes it evident that there are periods of zero counts in both series, for example from 1910 to 1920 for

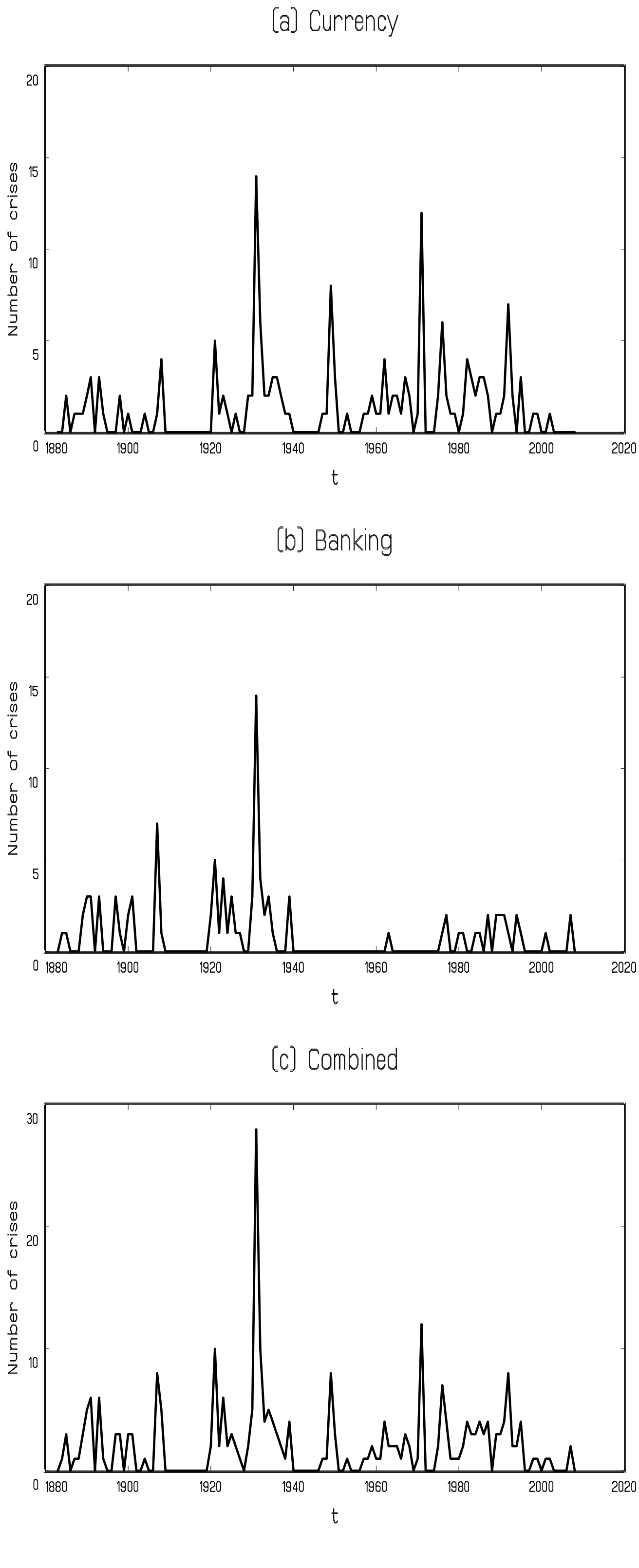

Figure 1: Number of countries experiencing a currency or a banking crisis, per annum, 1883 to 2008.

Table 1:

Number of currency and banking crises per country, 1883 to 2008.

Country	Currency	Banking	Combined
Argentina	20	9	4
Australia	7	2	0
Belgium	5	5	0
Brazil	14	9	3
Canada	10	1	0
Chile	10	5	1
Denmark	8	6	2
Finland	7	5	3
France	9	6	0
Germany	5	3	1
Greece	7	1	1
Italy	8	8	1
Japan	7	4	0
Netherlands	6	3	1
Norway	4	5	1
Portugal	6	5	2
Spain	8	5	1
Sweden	5	5	2
Switzerland	4	2	0
UK	11	2	0
US	7	10	2
Total	168	101	25

currency crises and from 1940 to 1960 for banking crises. This may suggest the need for more than a single regime to be used for modelling these data, a feature seen to be borne out below.

Table 2 provides some descriptive statistics of the data on currency and banking crises. For the currency, banking and combined number of crises, the data exhibit some over-dispersion with the sample variance approximately three times as large as the sample mean for the currency and banking crises, and approximately 5 times as large for the combined number of crises. Also given in the table are the total number of crises over the sample period. Of the 126 years analysed from 1883 to 2008, in nearly 55% of these years there is at least one country experiencing a currency crisis, and just over 34% experiencing a banking crisis.

Table 2:

Descriptive statistics on the number of countries experiencing either a currency or banking crisis, or both, 1883 to 2008.

Statistic	Currency	Banking	Combined
Mean	1.333	0.802	2.135
Variance	4.560	2.912	11.078
Number of years where a crisis occurs			
0	57	83	47
1	30	18	23
2	19	11	16
3	10	9	15
4	3	2	10
5	1	1	4
6	2	0	3
7	1	1	1
8	1	0	3
9	0	0	0
10	0	0	2
11	0	0	0
12	1	0	1
13	0	0	0
14	1	1	0
15	0	0	0
16	0	0	0
17	0	0	0
18	0	0	0
19	0	0	0
20	0	0	0
21	0	0	0
22	0	0	0
23	0	0	0
24	0	0	0
25	0	0	0
26	0	0	0
27	0	0	0
28	0	0	1
29	0	0	0
30	0	0	0
	126	126	126

Table 3:

Estimated ACF and PACF of currency or banking crises data.

Lag	Currency		Banking		Combined	
(1 year)	ACF	PACF	ACF	PACF	ACF	PACF
1	0.235	0.235	0.306	0.306	0.310	0.310
2	0.027	-0.030	0.121	0.031	0.085	-0.012
3	0.037	0.040	0.083	0.042	0.040	0.019
4	0.026	0.009	0.076	0.040	0.012	-0.005
5	0.088	0.085	-0.022	-0.068	0.013	0.011
6	0.015	-0.028	0.081	0.108	0.002	-0.006
7	-0.038	-0.036	0.075	0.026	-0.035	-0.038
8	-0.010	0.003	0.187	0.166	0.074	0.106
9	0.015	0.016	0.043	-0.073	-0.005	-0.063
10	0.065	0.057	0.184	0.182	0.105	0.136
11	-0.018	-0.049	0.032	-0.098	-0.044	-0.137
12	-0.064	-0.045	-0.076	-0.105	-0.135	-0.095
13	-0.003	0.021	-0.048	0.015	-0.082	-0.012
14	0.001	-0.006	0.040	0.016	-0.030	0.007
15	-0.005	-0.012	-0.124	-0.135	-0.085	-0.073
16	-0.010	0.000	0.001	0.041	-0.048	-0.012
17	-0.010	0.007	-0.062	-0.087	-0.089	-0.050
18	0.109	0.116	-0.019	-0.022	0.046	0.075
19	-0.065	-0.136	-0.120	-0.078	-0.107	-0.155
20	-0.154	-0.119	-0.095	-0.055	-0.190	-0.135

As a preliminary analysis of the dynamic structure of the data, Table 3 contains the sample autocorrelation function (ACF) and the sample partial autocorrelation function (PACF) for the crisis data. The ACF and the PACF each show strong evidence of first order serial dependence for both the currency and banking crises series. Also present are longer term dynamics, especially in the case of banking crises, where there is evidence of a potential ten year cycle. Though the serial correlations are not particularly strong, the episodic nature of both series suggests it may be prudent to entertain both autoregressive and moving average dynamics and, possibly, regime switching.

3 A Threshold INARMA Model of Crises

The nature of the crisis data is that there is little or no activity in the series for prolonged periods of time, that is runs of counts which may be zero or one, followed by relatively short bursts of major crisis activity; these bursts of activity, by casual observation, often seem to appear at around 10 year intervals. The data also exhibit overdispersion whereby the volatility of countries in crisis mode tends to be larger the average number.

To capture the empirical features of currency and banking crises as identified in Section 2, let y_t represent the number of countries that are in crisis at time t. To capture the transmission of crises across countries, consider the INAR(1) model (for a recent review of integer time series models, see Jung and Tremayne (2006b))

$$y_t = \alpha_1 \circ y_{t-1} + u_t, \qquad t = 1, 2, \cdots T, \tag{1}$$

where the parameter α_1 controls the strength of autocorrelation in y_t between adjacent time periods and $\alpha_1 \circ y_{t-1}$ is the "thinning" operator controlling the number of countries that are no longer in crisis mode. Formally, this operator is defined as (Al-Osh and Alzaid (1987, 1988), Alzaid and Al-Osh (1990))

$$\alpha_1 \circ y_{t-1} = \sum_{s=1}^{y_{t-1}} e_{s,t-1}, \tag{2}$$

where $e_{s,t-1}$ is a Bernoulli random variable with probability α_1 of a country remaining in a crisis the next period. The thinning operator has the property of preserving the integer status of the random variable y_t in contrast to the usual continuous time operator of the AR(1) model for example. In the extreme case where all y_{t-1} countries that are in crisis mode, stop being in crisis, all draws of the Bernoulli random variable would equal zero. Assuming independence the probability of this occurring is $(1 - \alpha_1)^{y_{t-1}}$. The random variable u_t allows new countries to go into crisis mode and hence add to the existing number of crisis affected countries. More formally, $\{u_t\}$ is a sequence of *iid* random shocks with constant mean λ. The higher is λ, the higher is the expected number of countries going into a financial crisis. In the case where these shocks are equidispersed, a Poisson distribution is commonly specified for u_t. Alternatively, in the case of overdispersion where the volatility of shocks is greater than λ, a negative binomial distribution can be specified. To allow for additional factors which can impact upon the transmission of crises, define λ_t which is expressed as a function of a set of regressors. This set may include dummy variables to account for the effects of large shock events such as the Great Crash in the early 1930s, or exogenous factors which can speed up or even slow down, the transmission channel.

The INAR(1) model in equation (1) can capture many of the features of financial crises, whereby there are periods where countries enter a crisis, so y_t is positive, and periods where there is tranquility where y_t is either zero or a small number, say $y_t = 1$. However, a feature of the data discussed in Section 2 is that there are relatively long periods where countries do not necessarily experience a financial crisis. One way to proceed is to distinguish between two periods where one period represents a noncrisis period and another representing a crisis period. This strategy is also commonly adopted in the contagion literature where the effects of additional shocks across countries and financial markets that occur only during financial crises are modelled by distinguishing between periods of crisis and noncrisis periods (Dungey, Fry, González-Hermosillo and Martin (2010)). Formally this is modelled by assuming that two regimes operate and that these are triggered by a threshold depending upon the value of y_{t-1}. Thresholds are relatively easy to model using integer data because the support for the threshold is not the real line as it is with continuous models. In the case of the crisis data, be it for banking or currency, the threshold would likely be 1, or, possibly, 2. It is not difficult to use hypothesis testing methods to confirm a preferred value. Suppose the threshold indicates regime 1 occurs for y_t if $y_{t-1} \leq j$ and regime 2 for $y_{t-1} > j$; the threshold is designed to dichotomize the sample, but for pragmatic estimation reasons it is necessary that there is at least a 'reasonable' number of observations in each regime. Extending the INAR(1) in equation (1) to allow for thresholds leads to the threshold INAR model, denoted as TINAR(1), of the form

$$y_t = \alpha_1 \circ y_{t-1} + (\gamma_1 \circ y_{t-1}) d_{y_{t-1} > j} + u_t, \tag{3}$$

where $d_{y_t > j}$ is an indicator variable which is unity if the condition is true, and zero otherwise. Following Du and Li (1991), the two thinning operations in (3), $\alpha_1 \circ y_{t-1}$ and $\gamma_1 \circ y_{t-1}$, are treated independently. The probability of countries remaining in a crisis during a financial crisis period is $\alpha_1 + \gamma_1$, while now α_1 represents the corresponding probability in a noncrisis period. As financial crises by definition occur where the number of countries going into crisis mode and remaining in this state is widespread, this would suggest that $\gamma_1 > 0$, resulting in an increase in autocorrelation in periods of financial crises.

A further feature of the crisis data discussed in Section 2 is the episodic nature of financial crises suggesting a renewal of crisis activity of approximately every 10 years. To capture this feature a 10th lag term is also included in the crisis regime. A prototype TINAR(10) crisis model is

$$y_t = \alpha_1 \circ y_{t-1} + (\gamma_1 \circ y_{t-1} + \gamma_{10} \circ y_{t-10}) d_{y_{t-1}>j} + u_t. \tag{4}$$

One might also make the two regimes in equation (4) even more separate by having different entry processes, because there are more entries during times of crisis than not. Again, this is effectively a regime dummy variable regressor and is trivial to incorporate. But it can also be appreciated that, after a noncrisis period, major crisis activity can only be modelled with a TINAR-type structure by entries to the system through the random variables u_t. Moreover, a crisis affected country may remain in the system for a few years before the crisis period dies down again and normal activity is resumed for that country. To emphasise this feature, it may be fruitful to add a moving average component to the model, as adopted by Martin, Tremayne and Jung (2012). Extending their approach to threshold integer models, a TINARMA(10,1) model is specified as

$$\begin{aligned} x_t &= \alpha_1 \circ x_{t-1} + (\gamma_i \circ x_{t-1} + \gamma_{10} \circ x_{t-10})\, d_{y_t>j} + u_t \\ y_t &= x_{t-1} + \beta_1 \circ u_t, \end{aligned} \tag{5}$$

where y_t is the observed count variable, and x_t is a latent (unobserved) count variable. The moving average term is controlled by the parameter β_1. This form of the TINARMA model is based on the specification proposed by McKenzie (1988) which uses the reversibility property of the moving average term. In the special case where $\beta_1 = 0$, equation (5) reduces to the TINAR(10) specification of equation (4). Of course further extensions can be entertained by including moving average terms in the crisis regime.

To gain some insight into the properties of the TINARMA model as a model of financial crises, Figure 2 presents simulated time series data of a TINARMA(1,0) model with negative binomial innovations to capture overdispersion. The model is given by

$$\begin{aligned} y_t &= \alpha_1 \circ y_{t-1} + (\gamma_1 \circ y_{t-1}) d_{y_{t-1}>j} + u_t, \\ u_t &\sim NB(\lambda, \omega), \end{aligned} \tag{6}$$

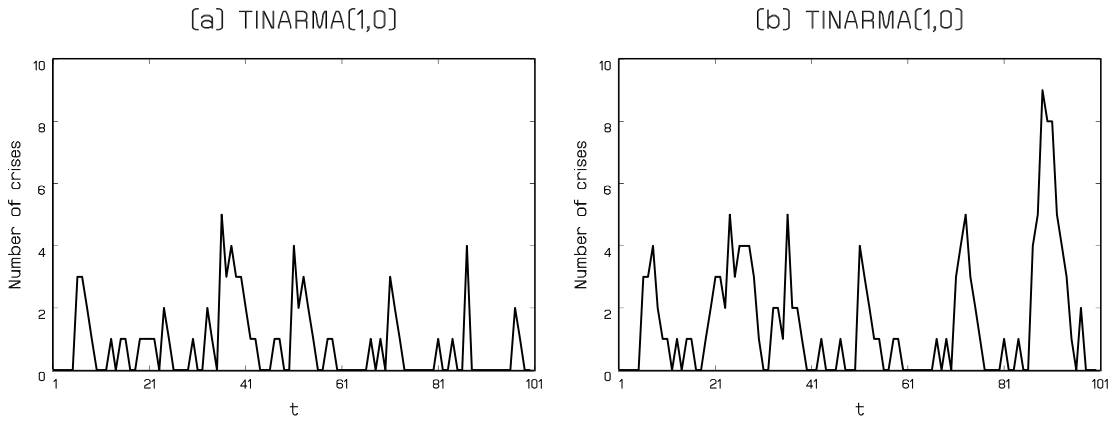

Figure 2: Simulated crises based on the TINARMA(1,0) model: effect of increasing autocorrelation during the crisis period from (a) $\gamma_1 = 0.5$ to (b) $\gamma_1 = 0.7$.

where the parameters for Figure 2(a) are $\theta = \{\alpha_1 = 0.1, \gamma_1 = 0.5, \lambda = 0.4, \omega = 0.5\}$, and for Figure 2(b) are $\theta = \{\alpha_1 = 0.1, \gamma_1 = 0.7, \lambda = 0.4, \omega = 0.5\}$. Figure 2(a) shows that the noncrisis period exhibits little evidence of temporal dependence whereas during the crisis period there is a jump in the number of crises as well as the duration of the crises. The effect of increasing the strength of this dependence during the crisis period is highlighted in Figure 2(b) where both the number and the duration of crises increases which occurs by increasing the crisis autocorrletion parameter from $\gamma_1 = 0.5$ to $\gamma_1 = 0.7$.

The effect of increasing the overdispersion parameter in the INARMA(1,0) model in equation (6) from $\omega = 0.5$ to $\omega = 0.6$, is demonstrated in Figure 3, with the parameters set for Figure 3(a) at $\theta = \{\alpha_1 = 0.1, \gamma_1 = 0.5, \lambda = 0.4, \omega = 0.5\}$, and for Figure 2(b) at $\theta = \{\alpha_1 = 0.1, \gamma_1 = 0.5, \lambda = 0.4, \omega = 0.6\}$. As an increase in the overdispersion parameter increases the variance at a faster rate than the mean of the innovation distribution, the number of financial crises increases in both size and duration.

4 Econometric Methods

This section provides a discussion of the estimation and inferential methods applied to the TINARMA model. In both cases simulation methods are employed. In the case of estimation, the approach is based on an efficient method of moments estimator

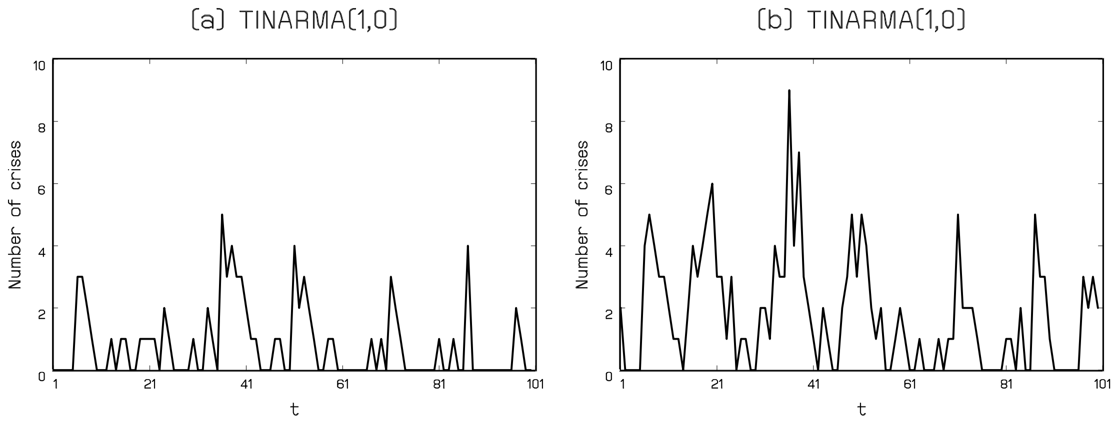

Figure 3: Simulated crises based on the TINARMA(1,0) model: effect of increasing overdispersion during the crisis period from (a) $\omega_1 = 0.5$ to (b) $\omega_1 = 0.6$.

(EMM) as estimation by maximum likelihood is not feasible for this general class of models. In the case of inference, simulation methods based on subsampling are adopted to circumvent the computational problems in generating standard errors of the EMM estimator for the TINARMA class of models.

4.1 Estimation

Whilst it is straightforward to specify a general integer time series model that allows for arbitrary lag structures with dynamics characterised by both autoregressive and moving average components, estimation of this class of models has proven to be problematic. Maximum likelihood is relatively straightforward for the INAR(1) model as it is possible to derive the log-likelihood function and estimate the unknown parameters using a standard gradient algorithm (Al-Osh and Alzaid (1987)). For higher order INARMA models, the log-likelihood becomes increasingly more complicated, but nonetheless can be estimated in theory at least (Bu, Hadri and McCabe (2008)). For INMA or mixed INARMA models, maximum likelihood is no longer feasible as it is now not even possible to write out the log-likelihood function for this class of models (see Jung, Ronning and Tremayne (2005) for a discussion of alternative estimators). To circumvent this problem, Martin, Tremayne and Jung (2012) propose a simulation based approach using an efficient method of moments estimator (EMM). The motivate behind this strategy is that even though INARMA models are difficult, if not impossible, to estimate using standard econometric methods based on maximum likelihood, these models are relatively easy to simulate. Moreover, in implementing the EMM estimator it is necessary to be able to choose a model that provides a suitable approximation of the true log-likelihood, commonly referred to as the auxiliary model. For this class of models the continuous time ARMA class of models represents a suitable choice. Even though this class of models ignores the integer status of the dependent variable being analysed, nonetheless consistent and asymptotically efficient parameter estimates are obtained under certain regularity conditions using EMM by calibrating the simulated data with the actual data via the auxiliary model (Gallant and Tauchen (1996); Gouriéroux, Monfort, and Renault (1993); Gouriéroux and Monfort (1994) Duffie and Singleton (1993) and Smith (1993)). The finite sample performance of the EMM estimator is investigated by Martin, Tremayne and Jung (2012) in the case of the INARMA model using a range of Monte Carlo experiments which is compared to the maximum likelihood estimator of Al-Osh and Alzaid (1987) and the constrained least squares estimator of Klimko and Nelson (1978), when either of these two estimators are feasible.

Consider the TINARMA$(p,1)$ class of models

$$\begin{aligned} x_t &= \sum_{i=1}^{p} \alpha_i \circ x_{t-i} + \sum_{i=1}^{p} \left(\gamma_i \circ x_{t-i}\right) d_{y_t>j} + u_t \\ y_t &= x_{t-1} + \beta_1 \circ u_t \\ u_t &\sim NB\left(\lambda, \omega\right), \end{aligned} \tag{7}$$

where y_t is the observed count variable, x_t is a latent (unobserved) count variable, $d_{y_t>j}$ is the indicator variable and the $2p+1$ thinning operations in (7) are all treated independently as before. The variable $d_{y_t>j}$ is an indicator variable which is unity if the condition is true, and zero otherwise. The choice of the threshold number of countries in crises is represented by j, can be determined empirically or chosen based on a definition of what constitutes a crisis situation. Let the unknown parameters be given by

$$\theta = \left\{\alpha_1, \alpha_2, \cdots, \alpha_p, \gamma_1, \gamma_2, \cdots, \gamma_p, \lambda, \omega, \beta_1\right\}. \tag{8}$$

The steps involved to implement the EMM estimator for this class of models are as follows. First, choose a given set of starting parameter values of the TINARMA(p,1) model given by $\theta_{(0)}$. Second, simulate (7) for a given choice of $\theta_{(0)}$ by expressing the thinning operations in terms of independent uniform random numbers $e^{(i)}_{s,t-i}$, $w^{(i)}_{s,t-i}$, $h_{s,t}$ $i = 1, 2, \cdots, p$, and $h_{s,t}$ according to

$$\begin{aligned}
x_{s,t} &= \sum_{i=1}^{p}\sum_{s=1}^{x_{t-i}} e^{(i)}_{s,t-i} + \sum_{i=1}^{p}\sum_{s=1}^{x_{t-i}} w^{(i)}_{s,t-i} d_{y_{s,t}>j} + u_t \\
y_{s,t} &= x_{s,t-1} + \sum_{s=1}^{u_t} h_{s,t} \qquad (9)\\
u_t &\sim NB(\lambda, \omega),
\end{aligned}$$

where the uniform random numbers have moments

$$\begin{aligned}
E\left[e^{(i)}_{s,t-i}\right] &= \alpha_i, \quad i = 1, 2, \cdots, p, \\
E\left[w^{(i)}_{s,t-i}\right] &= \gamma_i, \quad i = 1, 2, \cdots, p, \qquad (10)\\
E\left[h_{s,t}\right] &= \beta_1.
\end{aligned}$$

The TINARMA model is simulated for a period of length N, where N is commonly chosen as a factor of the length of the time series data, T. In the empirical analysis N is set at $N = 100T$.

Third, specify the auxiliary model

$$y_t = \phi_0 + \sum_{i=1}^{k} \phi_i y_{t-i} + \sum_{i=1}^{k} \delta_i y_{t-i} d_{y_{t-1}>j} + v_t, \qquad (11)$$

which is a threshold AR(k) model with a constant term where v_t is *iid* $N(0, \sigma_v^2)$. The choice of k needs to be large enough to be able to identify the p autoregressive parameters for each regime in equation (7) and the moving average parameter β_1 from higher order lags in the auxiliary model specified in (11). This choice is also motivated by the property that a finite moving average model can be represented by an infinite autoregressive model. In practice the choice of the lag length in (11) is finite, with the quality of the approximation improving as the lag length is increased. Martin, Tremayne and Jung (2012) find that for values of $k \leq 3$ in INARMA models, provides a suitable approximation to estimate moving average parameters. The $(2k + 2)$ gradient vector of the auxiliary model at time t is

$$g_t = \left\{ \frac{v_t}{\sigma_v^2}, \frac{v_t y_{t-1}}{\sigma_v^2}, \cdots, \frac{v_t y_{t-k}}{\sigma_v^2}, \frac{v_t y_{t-1} d_{y_{t-1}>j}}{\sigma_v^2}, \cdots, \frac{v_t y_{t-k} d_{y_{t-1}>j}}{\sigma_v^2}, \left(\frac{v_t^2}{\sigma_v^2} - 1\right)\frac{1}{2\sigma_v^2} \right\}, \qquad (12)$$

where from (11) $v_t = y_t - \phi_0 - \sum_{i=1}^{k} \phi_i y_{t-i} - \sum_{i=1}^{k} \delta_i y_{t-i} d_{y_{t-1}>j}$. The choice of the first and last moment conditions in (12) are used to identify the parameters λ and ω of the negative binomial distribution. This property of the moment condition arises because the mean of the negative binomial distribution is $\lambda\omega/(1-\omega)$, which is captured by the intercept term in the auxiliary model and the variance of the negative binomial distribution is $\lambda\omega/(1-\omega)^2$, which is captured by σ_v^2. This suggests that the following 2 equations

$$\phi_0 = \frac{\lambda\omega}{1-\omega}, \qquad \sigma_v^2 = \frac{\lambda\omega}{(1-\omega)^2},$$

can be rearranged to identify λ and ω as

$$\omega = 1 - \frac{\phi_0}{\sigma_v^2}, \qquad \lambda = \frac{\phi_0^2}{\sigma_v^2 - \phi_0}.$$

Fourth, the auxiliary model in (11) is estimated by least squares using the actual data (y_t) with parameter estimates given by

$$\widehat{\psi} = (\widehat{\phi}_0, \widehat{\phi}_1, \cdots, \widehat{\phi}_k, \widehat{\delta}_1, \widehat{\delta}_2, \cdots, \widehat{\delta}_k, \widehat{\sigma}_v^2)', \tag{13}$$

and $\widehat{\sigma}_v^2 = T^{-1}\sum_{t=1}^{T}(y_t - \widehat{\phi}_0 - \sum_{i=1}^{k}\widehat{\phi}_i y_{t-i} - \sum_{i=1}^{k}\widehat{\delta}_i y_{t-i} D_{y_{t-1}>j})$, the variance of the least squares residuals associated with equation (11). It is assumed that the data length is $T+k$ to allow for conditioning on the first k observations. These estimates of the auxiliary model are also equivalent to the constrained least squares estimates of Klimko and Nelson (1978).

Finally, the EMM estimator is based on solving

$$\widehat{\theta} = \arg\min_{\theta} G_s' I^{-1} G_s = \arg\min_{\theta} Q(\theta), \tag{14}$$

using an iterative algorithm. The $(2k+2)$ vector G_s represents the gradients of the auxiliary model in (12) with y_t replaced by the simulated data $y_{s,t}$, and the auxiliary parameters are evaluated at $\widehat{\psi}$ in (13)

$$G_s = \frac{1}{N}\sum_{i=1}^{N} g_{s,t}, \tag{15}$$

where

$$g_{s,t} = \left\{ \frac{v_{s,t}}{\widehat{\sigma}_v^2}, \frac{v_{s,t}y_{s,t-1}}{\widehat{\sigma}_v^2} \cdots \frac{v_{s,t}y_{s,t-k}}{\widehat{\sigma}_v^2}, \frac{v_{s,t}y_{s,t-1}d_{y_{s,t-1}>j}}{\widehat{\sigma}_v^2} \cdots \frac{v_{s,t}y_{s,t-k}d_{y_{s,t-1}>j}}{\widehat{\sigma}_v^2}, \left(\frac{v_{s,t}^2}{\widehat{\sigma}_v^2} - 1\right)\frac{1}{2\widehat{\sigma}_v^2} \right\}, \tag{16}$$

and

$$\widehat{v}_{s,t} = y_{s,t} - \widehat{\phi}_0 - \sum_{i=1}^{k}\widehat{\phi}_i y_{s,t-i} - \sum_{i=1}^{k}\widehat{\delta}_i y_{s,t-i} d_{y_{s,t-1}>j}.$$

The $(2k+2) \times (2k+2)$ matrix I in (14) is computed as (Gallant and Tauchen (1996))

$$I = \frac{1}{T}\sum_{t=1}^{T} g_t g_t', \tag{17}$$

which is the outer product of the gradients of the auxiliary model evaluated using the actual data.

By construction, evaluating the gradient vector G_s in (14) at the actual data y_t and not the simulated data $y_{s,t}$, produces a vector of zeros. This suggests that the EMM estimator chooses parameter values of θ when the simulated data from the TINARMA model match the actual data which occurs by minimizing $Q(\theta)$. For a just-identified model $Q(\widehat{\theta}) = 0$ as there is an equal number of equations to set all gradients to zero. For an over-identified model $Q(\widehat{\theta}) > 0$, but provided that the model is correctly specified the value of the objective function should not be statistically significant from zero. Formally, an overall test of the model's specification is based on Hansen's J-test given by

$$J = TQ(\widehat{\theta}),$$

which is distributed asymptotically under the null hypothesis as χ^2_r, where r is the number of over-identifying restrictions, equal to $\dim(\psi) - \dim(\theta)$.

As a result of the nonlinear structure of the model analytical solutions of the parameter estimates are not available, making it necessary to use numerical procedures. The approach adopted here is to use a grid search method as standard gradient algorithms can break down as the numerical gradients may not always be differentiable at all parameter values as a result of the integer status of the dependent variable. Martin, Tremayne and Jung (2012) find that this algorithm performs very well for a broad range of INARMA specifications. Some Monte Carlo evidence is given next for the TINARMA model.

4.2 Monte Carlo Evidence

To gain some insight into the finite sample properties of the EMM estimator for the TINARMA model, the results of some Monte Carlo experiments are now reported. The first experiment is based on the TINARMA(1,0) model with negative binomial innovations

$$\begin{aligned} y_t &= \alpha_1 \circ y_{t-1} + (\gamma_1 \circ y_{t-1}) d_{y_{t-1}>1} + v_t \\ v_t &\sim NB(\lambda, \omega), \end{aligned} \tag{18}$$

where the thinning operators in (18) are assumed to be independent. The true parameters are set at

$$\theta = \{\alpha_1 = 0.1, \gamma_1 = 0.6, \lambda = 0.4, \omega = 0.5\}.$$

The EMM estimator is based on simulation runs of length $N = 100T$, where the sample sizes are chosen as $T = \{100, 200\}$ which nest the sample size of the data used in the empirical section. The auxiliary model is based on the threshold AR model given in (11) where the lag structure varies from $k = 1$ to $k = 3$. In solving (14) the number of random searches is restricted to 200. All calculations are performed using the software GAUSS Version 10, while the random number generators are **rndlcnb** for the negative binomial distribution and **rndlcu** for the uniform distribution, where the latter is used in computing the thinning operations.

Table 4 provides summary statistics on the finite sample distributions of the EMM estimator based on 2000 replications. The estimate of the noncrisis autoregressive parameter α_1, is biased upwards while the opposite is true for the crisis autoregressive parameter γ_1, which is biased downwards. The results for the variance parameter estimates show that the estimate of λ is biased upwards whereas the estimate of ω exhibits downward biasedness. For all parameter estimates the size of the biasedness and the RMSE decrease as the sample size is increased from $T = 100$ to $T = 200$. There results in general are qualitatively the same as the Monte Carlo results reported in Martin, Tremayne and Jung (2012) who find that for (non-threshold) INARMA models with Poisson innovations the EMM estimator, together with other estimators including maximum likelihood and the conditional least squares estimator of Klimko and Nelson (1978), that the autoregressive parameters are biased downwards while the variance parameter is biased upwards. A comparison of the mean and RMSE sample statistics across different lag structures used in the auxiliary model of the EMM estimator suggests that the just-identified case of a lag of $k = 1$ tends to work best in terms of minimizing the finite sample bias and the RMSE.

In the second experiment the DGP in equation (18) is extended to allow for an MA component, resulting in the following TINARMA(1,1) model with negative binomial

Table 4:
Monte Carlo results for the EMM estimator of the TINARMA(1,0) model. Sample means based on 2000 replications with RMSE given in parentheses. The population parameters are $\theta = \{\alpha_1 = 0.1, \gamma_1 = 0.6, \lambda = 0.4, \omega = 0.5\}$.

Aux. Model Lags	$\widehat{\alpha}_1$	$\widehat{\gamma}_1$	$\widehat{\lambda}$	$\widehat{\omega}$
		$T = 100$		
$k = 1$	0.157	0.439	0.558	0.403
	(0.146)	(0.252)	(0.275)	(0.153)
$k = 2$	0.170	0.436	0.573	0.368
	(0.165)	(0.262)	(0.291)	(0.183)
$k = 3$	0.181	0.423	0.590	0.339
	(0.178)	(0.277)	(0.306)	(0.205)
		$T = 200$		
$k = 1$	0.140	0.501	0.515	0.435
	(0.116)	(0.186)	(0.227)	(0.119)
$k = 2$	0.140	0.504	0.527	0.411
	(0.116)	(0.190)	(0.242)	(0.139)
$k = 3$	0.156	0.489	0.553	0.386
	(0.133)	(0.206)	(0.267)	(0.162)

innovations

$$\begin{aligned} x_t &= \alpha_1 \circ x_{t-1} + (\gamma_1 \circ x_{t-1})\, d_{y_t > 1} + u_t \\ y_t &= x_{t-1} + \beta_1 \circ u_t \\ u_t &\sim NB(\lambda, \omega), \end{aligned} \tag{19}$$

The true parameters are set at

$$\theta = \{\alpha_1 = 0.1, \gamma_1 = 0.3, \beta_1 = 0.3, \lambda = 0.4, \omega = 0.5\}.$$

In choosing the lag structure of the auxiliary model used to compute the EMM estimates, the smallest lag length is of order $k = 2$ to be able to identify both the autoregressive (α_1) and moving average (β_1) parameters in the crisis period.

The finite sample properties of the EMM estimator based on the TINARMA(1,1) model in equation (19) are given in Table 5. The results in the case of the autoregressive parameters, α_1 and γ_1, and the variance parameters λ and ω, are qualitatively the same as they are for the INARMA(1,0) model using equation (18) as the data generating

Table 5:

Monte Carlo results for the EMM estimator of the TINARMA(1,1) model. Sample means based on 2000 replications with RMSE given in parentheses. The population parameters are $\theta = \{\alpha_1 = 0.1, \gamma_1 = 0.3, \beta_1 = 0.3, \lambda = 0.4, \omega = 0.5\}$.

Aux. Model Lags	$\widehat{\alpha}_1$	$\widehat{\gamma}_1$	$\widehat{\beta}_1$	$\widehat{\lambda}$	$\widehat{\omega}$
			$T = 100$		
$k = 2$	0.143	0.428	0.309	0.606	0.343
	(0.122)	(0.278)	(0.207)	(0.306)	(0.201)
$k = 3$	0.154	0.439	0.318	0.629	0.307
	(0.136)	(0.283)	(0.221)	(0.324)	(0.228)
			$T = 200$		
$k = 2$	0.132	0.406	0.300	0.563	0.390
	(0.095)	(0.252)	(0.162)	(0.266)	(0.154)
$k = 3$	0.144	0.429	0.286	0.598	0.357
	(0.110)	(0.272)	(0.171)	(0.292)	(0.181)

process. Focussing on the moving average parameter β_1, the results show that the EMM estimator exhibits little biasedness showing that this parameter is well resolved. A comparison of the finite sample results for the case where the lag structure of the auxiliary model is $k = 2$ and $k = 3$, shows that the performance of the EMM estimator is best for the smaller lag structure.

4.3 Inference

An attractive feature of the EMM estimator in the present context is that it does not require the specification of a likelihood function. This would be based on the conditional distribution of an observation, conditional on its past and is not available whenever moving average components are included in a thinning model. Hence, routine evaluation of appropriate estimated standard errors is not possible. Following Martin, Tremayne and Jung (2012), who encounter a similar difficulty, the subsampling approach of Politis, Romano and Wolf (1999) (PRW) is adopted. An important advantage of subsampling is that it is possible to derive standard errors in certain situations where a bootstrap approach would be invalid, whilst the requirements that must be met for it to provide a valid inference tool are mild, often requiring little more than stationarity of the underlying data generating process. An abbreviated account of the approach of Martin, Tremayne and Jung (2012) to estimate standard errors is presented here.

Suppose that $y_1, y_2, \cdots, y_T$ represents a sample of T observations on a stationary time series. Usually, statistical quantities are calculated using all the observations, apart, possibly, from end effects, whereas subsampling methods are based on making repeated computations of similar statistics based on subsamples of length B using the observations $y_i, y_{i+1}, \cdots, y_{i+B-1}$. To introduce relevant notation, let there be $j = 1, 2, \cdots, N_B$ such blocks used and suppose that a (scalar) parameter θ is estimated by $\widehat{\theta}_T$ using the full sample and by $\widehat{\theta}_{T,i,B}$ using the block of length B beginning at observation i. Provided $B \to \infty$ with T and $B/T \to 0$, the distribution of $\sqrt{B}(\widehat{\theta}_{T,i,B} - \theta)$ is Gaussian involving an unknown long-run variance that must be estimated and a suitable estimator of the variance of $\widehat{\theta}_T$, given by

$$\widehat{Var}_{T,B}(\widehat{\theta}_T) = \frac{B}{TN_B} \sum_{j=1}^{N_B} (\widehat{\theta}_{T,B,j} - \widehat{\theta}_{T,B,.})^2, \tag{20}$$

where $\widehat{\theta}_{T,B,.} = N_B^{-1} \sum_{j=1}^{N_B} \widehat{\theta}_{T,B,j}$, compare PRW, equation (3.40).

Difficulties that remain to be addressed include: what block length B to use; how many blocks N_B to use; and how to estimate the unknown long-run variance σ^2_∞, which is also needed. PRW, Section 3.8.2 show that, in the case of a sample mean at least, it is preferable to use the maximum available number of overlapping blocks of size B, *viz.* $N_B = T - B + 1$. This is because it can be shown that the variance estimator based upon subsampling using all available blocks is 50% more efficient, asymptotically, than the Carlstein (1986) estimator (see equation (3.46) in PRW) and so the approach recommended by PRW is adopted here. There remains the crucial choice of block length B. Work relating to the sample mean indicates that B should be $O(T^{1/3})$ and that the asymptotic mean squared error of the estimated variance in this case is $T^{1/3}$ times a complex quantity depending on the long-run variance σ^2_∞ (see PRW equation (9.4)). For more general statistics such as the EMM parameter estimators under consideration in this paper, not much is known about the choice of B. However, developing arguments in Chapters 9 and 10 in PRW (see Martin, Tremayne and Jung, 2012, Section 3.2 for fuller details) we use a number of different block lengths (usually 3 or 4) to estimate the long-run variance by determining it as the estimated intercept in a regression of the estimated values found from (20) on inverse block size for various values of B. We combine estimates based on B in a range $kT^{1/3}$ to $8kT^{1/3}$ with k

about 2 and our experience leads us to believe that such choices of block length with the crisis are likely to provide reliable estimates for asymptotic standard errors.

An exercise lending weight to this assertion is reported in Martin, Tremayne and Jung (2012), where estimated asymptotic standard errors in an INAR(1) model are reliably estimated for a case where the true asymptotic result is known. The standard errors of the EMM estimates are computed using 4 block lengths $B = \{8, 16, 32, 64\}$ with $T = 100$, which is fairly close to the length of data realisation with the crisis data under scrutiny here ($T = 126$). For each block size B the maximum number of data subsamples is $T - B + 1$. The EMM estimate is computed for each subsample using 100 searches, which, in turn, is used to compute an estimate of the variance of $\widehat{\theta}_{100,B}$. In the case of $B = 8$, this amounts to computing the EMM estimates $100 - 8 + 1 = 93$ times with the EMM objective evaluated in each case 100 times in performing the search procedure to minimize this objective function. The long-run variance σ^2_∞ is estimated as the intercept from a regression of the estimated variances corresponding to the four block sizes, on a constant and the regressor $\{1/8, 1/16, 1/32, 1/64\}$. The standard errors of $\widehat{\theta}_T$ are computed as $\widehat{\sigma}_\infty/\sqrt{T}$ where $\widehat{\sigma}^2_\infty$ is the estimate of σ^2_∞.

5 Empirical Results

The following TINARMA(10, 1) model is initially specified to model the autocorrelation structure of the currency, banking and combined crises data presented in Figure 1

$$\begin{aligned} x_t &= \alpha_1 \circ x_{t-1} + \gamma_{10} \circ x_{t-10} d_{y_t > j} + u_t \\ y_t &= x_{t-1} + \beta_1 \circ u_t \\ u_t &\sim NB(\lambda, \omega). \end{aligned} \tag{21}$$

The EMM estimates are based on solving (14) using 100,000 searches. The lag structure of the auxiliary model is $k = 1$ in (11) while the number of simulation paths used in the EMM algorithm to simulate the time path of the number of crises is set at $H = 100$.

The EMM parameter estimates are given in Table 6 for the currency and banking crises as well as for the combined crises. Standard errors are given in parentheses based on subsampling. In computing the standard errors the block lengths chosen are $\{16, 32, 64\}$ while the number of searches carried out for each subsample is set at 100. The results are reported for two definitions of crises with $j = 1$ corresponding to a

Table 6:
EMM estimates of TINARMA models with negative binomial innovations for the crisis data, with bootstrap standard errors given in parentheses. Threshold values given by j, expressed in terms of years.

Parameter	Currency		Banking		Combined	
	$j=1$	$j=2$	$j=1$	$j=2$	$j=1$	$j=2$
α_1	0.009	0.078	0.005	0.009	0.148	0.006
	(0.184)	(0.151)	(0.128)	(0.095)	(0.175)	(0.211)
γ_{10}	0.602	0.063	0.756	0.948	0.804	0.367
	(0.170)	(0.194)	(0.191)	(0.206)	(0222)	(0.116)
β_1	0.282	0.172	0.199	0.137	0.095	0.258
	(0.189)	(0.172)	(0.178)	(0.152)	(0.249)	(0.218)
λ	0.354	0.463	0.161	0.209	0.411	0.438
	(0.423)	(0.440)	(0.423)	(0.428)	(0.393)	(0.374)
ω	0.693	0.658	0.718	0.694	0.710	0.739

financial crisis occurring when there is more than one country in a crisis and with $j = 2$ corresponding to a financial crisis occurring when there is more than two countries in a crisis. The magnitude of the parameter estimates are consistent across the currency and banking crises with the noncrisis period demonstrating little autocorrelation structure with the strength of the dependence increasing during financial crises. For the noncrisis regime the autocorrelation structure is dominated by the first order moving average term suggesting that the memory of the process during this regime is around one year, although this parameter estimate is not statistically significant. For the crisis regime the results show strong evidence of a 10-year cycle, as identified in Section 2, in both the currency and banking crises.[2] Finally, the estimates of ω show strong evidence of over-dispersion.

6 Conclusions

The aim of this paper has been to construct an empirical model of crisis transmission across countries. Both banking and currency crises were investigated using annual data

[2]The model was also estimated where the length of the cycle varied for lags greater than or less than ten years. None of these alternative models improved upon the estimated model with a lag of 10 years.

on 21 countries over the period 1883 to 2008. An important feature of the data was that it was charcterised by low order counts corresponding to the number of countries in a crisis at a point in time. To capture these features of the data an integer ARMA model was used to model the dynamic transmission channels of financial crises. Moreover, as it was necessary to distinguish between periods of tranquility and periods of crises, a new class of models was developed called threshold integer ARMA models, or TINARMA. Estimation of the dynamic, low-order count model was based on Monte Carlo methods using efficient method of moments (EMM) as maximum likelihood methods were not feasible. The finite sample properties of this estimator were investigated using a range of Monte Carlo experiments. Monte Carlo methods were also used to construct standard errors of the parameter estimates using subsampling as the objective function corresponding to the EMM estimator is not differentiable everywhere for discrete time series models based on a finite simulation run.

The empirical results showed little evidence of autocorrelation during noncrisis periods, but strong evidence during periods of financial crises. An interesting result was the presence of a sporadic 10-year cycle in the data during periods of financial crises which was captured with the threshold model. The results also provided evidence of overdisperision with the volatility of crises exceeding that of the conditional mean.

In modelling the banking and currency crises the approach was univariate with the transmission of each crisis studied separately. This suggests that an important extension of this analysis is to allow for dynamic interactions between the two crises by performing a bivariate analysis. This would involve constructing a bivariate TINARMA model where crises would now be able to be transmitted across countries as well as across financial markets. To estimate this class of models a natural approach would be to specify a bivariate vector autoregression with threshold effects as the auxiliary model used in the computation of the EMM estimator. Having estimated a bivariate TINARMA model the dynamical interrelationships of crises could be revealed using impulse responses. As a result of the nonlinear features of the TINARMA model simulation based procedures could be used along the lines of generalized impulse response analysis (Koop, Pesaran, and Potter (1996)). A further extension of the empirical application is to investigate the predictive properties of the model by forecasting the presence of future crises. Again a natural approach to generate coherent forecasts of integer processes would be to use Monte Carlo methods to simulate the model into the future and map out probabilities associated with the number of countries in crisis at a particular point in time (Jung and Tremayne (2006a), McCabe and Martin (2005) and McCabe, Martin and Harris (2011)).

A Appendix: Data Construction

The binary data for the paper are taken from Dungey, Jacobs and Lestano (2010), who in turn collate data from published secondary sources. The sample comprises the 21 countries whose annual banking and currency crisis occurrences were collected in Bordo et al (2001); Argentina, Australia, Belgium, Brazil, Canada, Chile, Denmark, Finland, France, Germany, Greece, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, UK and the US. It is worth noting that a number of these economies have made a significant transition from developing to developed since the late 19th century.

The Bordo et al (2001) data begins in 1883 and terminates in 1998. Laeven and Valencia (2008) provide banking and currency crisis dates for 100 countries, including those in the Bordo et al (2001) data set. We utilise their data to update the Bordo et al (2001) data set to 2007. We provide an additional year of data by applying the Laeven and Valencia rule to identify currency crises in 2008 and detecting banking crises using information provided by the IMF on program applications.

A.1 Currency Crises

The Bordo et al (2001) currency crisis data are determined using an exchange market pressure index (see Eichengreen et al, 1996 for example), which consists of a weighted average of changes in the exchange rate against a numeraire, the interest rate differential against the numeraire interest rate and the relative change in reserve holdings compared with the numeraire country. The weights are given by the inverse volatility of each of these series to provide appropriate rescaling. The advantage of these exchange market pressure indices is that they detect pressure in floating, managed and fixed rate regimes, and are well understood in the international literature. A crisis is determined to occur when the exchange market pressure index exceeds some critical threshold; and the authors indicate they follow the earlier work of Eichengreen et al (1996) and set this threshold as the being the mean plus 3 standard deviations of the index. An unfortunate feature of this approach is that the threshold is sample specific, and sensitivity analysis typically reveals that different samples may result in different periods identified as crisis due particularly to the changing volatility observed in assets experiencing crisis conditions.

Laeven and Valencia (2008) nominate a simpler rule where currency crises for a country, i, are indicated by a 30 percent or greater nominal depreciation of their domestic currency over the previous year, with the additional condition that the rate of depreciation has also increased by 10 percent or more over the year before that - that is the rate of depreciation has been accelerating rapidly over the prior two year period. The same criteria are used for fixed exchange rate devaluations. We applied the Laeven and Valencia rule to update the dataset for 2008, and found no additional incidents of currency crises.

Multiple exceedances of the threshold criteria are a common problem in rules defining currency crises, and are generally resolved by the application of windows based on rules of thumb. This aspect seriously differentiates the binary data generated in crisis data sets from those generated in other literature, such as business cycle phases in Harding and Pagan (2006), who can write down the underlying data generating process more clearly. The windowing conventions applied in the current data set are to treat currency crises are indicated in the first year of a five year window in which they may occur in multiple instances - in Bordo et al (2001) this is justified by reference to the recovery period of GDP and in Laeven et al (2008) simply given as a rule of thumb.

A.2 Banking Crises

Banking crises are more difficult to discern than currency crises as there is generally less transparent data on the bank balance sheets. Bordo et al (2001) follow Caprio and Klingebiel (1997) who provide data from the 1970s onwards. They define bank crises as the situation where the ratio of nonperforming loans to total loans of 5 to 10 percent, arguing that this is a conservative measure of the cases where bank capital would be insufficient to cover the losses from these loans in the event that they were liquidated - that is the banking system is insolvent. Bordo et al adopt the Caprio and Klingebiel data where appropriate and use their method to provide earlier data.

Laeven and Valencia (2008) take a slightly different measure, where they consider the case of sharp increases in non-performing loans and exhaustion of capital in the banking system. They combine the numerical information with quantitative indicators to exclude episodes affecting only a small number of non-systemic institutions. We update the banking crisis data from the IMF based on the programs implemented in each of the countries in 2008.

References

[1] Al-Osh, M. and Alzaid, A.A. (1987), "First-order Integer-valued Autoregression (INAR(1)) Process", *Journal of Time Series Analysis*, 8, 261-275.

[2] Al-Osh, M. and Alzaid, A.A. (1988), "Integer-valued Moving Average (INMA) Process", *Statistical Papers*, 29, 281-300.

[3] Alzaid, A.A. and Al-Osh, M. (1990), "An Integer-valued pth Autoregressive Structure (INAR(p)) Process", *Journal of Applied Probability*, 27, 314-324.

[4] Bordo, M.J., Eichengreen, B. , Klingebield, D. and Martinez-Peria, M.S. (2001), "Is the Crisis Problem Growing More Severe?" *Economic Policy*, 16, 53-82.

[5] Brännäs, K. and Hall, A. (2001), "Estimation in Integer-valued Moving Average Models", *Applied Stochastic Models in Business and Industry*, 17,277-291.

[6] Brännäs, K. and Hellstrom, J. (2001), Generalized Integer-valued Autoregression, *Econometric Reviews*, 20, 425-443.

[7] Brännäs, K. and Shahiduzzaman Quoreshi, A.M.M. (2010), "Integer-valued Moving Average Modelling of the Number of Transactions in Stocks", *Applied Financial Economics*, 20, 1429-1440.

[8] Bu, R., Hadri, K. and McCabe, B.P.M. (2008), "Conditional Maximum Likelihood Estimation of Higher-Order Integer-valued Autoregressive Processes", *Journal of Time Series Analysis*, 29, 973-994.

[9] Carlstein, E., (1986), "The Use of Subseries Values for Estimating the Variance of a General Statistic from a Stationary Time Series", *Annals of Statistics*,14, 1171-1179.

[10] Caprio, G. and Klingebiel, D. (1997), "Bank Insolvencey: Bad Luck, Bad Policy or Bad Banking?", *Annual World Bank Conference on Development Economics 1996*, 1-26.

[11] Du, J.G and Li, Y. (1991), "The Integer-valued Autoregressive (INAR(p)) Model", *Journal of Time Series Analysis*, 12, 129-142.

[12] Duffie, D. and K. Singleton (1993), "Simulated Moments Estimator of Markov Models of Asset Prices", *Econometrica*, 61, 929-62.

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

[14] Dungey, M., Fry, R.A., González-Hermosillo, B. and Martin, V.L. (2010), *Transmission of Financial Crises and Contagion: A Latent Factor Approach*, Oxford University Press, Oxford.

[15] Dungey, M., Jacobs, J.P.A.M., and Lestano (2010), "The Internationalisation of Financial Crises: banking and currency crises 1883-2008", *manuscript*.

[16] Eichengreen, B., Rose, A.K. and Wyplosz, C. (1996), "Contagious Currency Crises: First Tests", *Scandinavian Journal of Economics*, 98, 463-484.

[17] Gallant A.R. and Tauchen, G. (1996), "Which Moments to Match?", *Econometric Theory*, 12 (4), 657-81.

[18] Gouriéroux, C. and Monfort, A. (1994), *Simulation Based Econometric Methods*, CORE Discussion Paper.

[19] Gouriéroux, C., Monfort, A. and Renault, E. (1993), "Indirect Inference", *Journal of Applied Econometrics*, 8, S85-S118.

[20] Jung, R.C., Ronning, G. and Tremayne, A.R. (2005), "Estimation in Conditional First Order Autoregression with Discrete Support", *Statistical Papers*, 46, 195-224.

[21] Jung, R.C. and Tremayne, A.R. (2006a), "Coherent Forecasting in Integer Time Series Models", *International Journal of Forecasting*, 22, 223-238.

[22] Jung, R.C and Tremayne, A.R. (2006b), "Binomial Thinning Models for Integer Time Series", *Statistical Modelling*, 6, 81-96.

[23] Klimko, L.A. and Nelson, P.I. (1978), "On Conditional Least Squares Estimation for Stochastic Processes", *Annals of Statistics*, 6, 629-642.

[24] Koop G., M.H. Pesaran, S.M. Potter (1996), "Impulse Response Analysis in Nonlinear Multivariate Models", *Journal of Econometrics,* 74, 119-147.

[25] Laeven, L. and Valencia, F. (2008), "Systemic Banking Crises: A New Database", IMF Working Paper, WP/08/224.

[26] McCabe, B.P.M. and Martin, G.M. (2005), "Bayesian Predictions of Low Count Time Series", *International Journal of Forecasting*, 21, 315-330.

[27] McCabe, B.P.M., Martin, G.M. and Harris, D. (2011), "Efficient Probabilistic Forecasts for Counts", *Journal of the Royal Statistical Society B*, 73, 253-272.

[28] McKenzie, E. (1988), "Some ARMA Models for Dependent Sequences of Poisson Counts", *Advances in Applied Probability*, 20,822-835.

[29] Politis, D.N., Romano, J.P. and Wolf, M. (1999), "Subsampling", New York: Springer-Verlag.

[30] Smith, A.A. (1993), "Estimating Nonlinear Time Series Models Using Simulated Vector Autoregressions", *Journal of Applied Econometrics*, 8, S63-S84.