Regionality Revisited: An Examination of the Direction of Spread of Currency Crises.

By Amil Dasgupta[1], Roberto Leon-Gonzalez[2] and Anja Shortland[2],

[1]London School of Economics
Houghton Street
London, WC2A 2AE
United Kingdom

and

[2]Department of Economics
University of Leicester
Leicester, LE1 7RH
United Kingdom

Preliminary. Comments welcome.

Abstract: We examine data on waves of currency crises from 1992, 1994, 1997, and 1998 to evaluate several hypotheses on the determinants of contagion. We simultaneously consider trade, financial links, legal and institutional indicators, geographical distance, and standard macroeconomic variables as controls. To overcome data limitations and account for model uncertainty, we utilize Bayesian methodologies hitherto unused in the empirical literature on contagion. In particular, we use the Bayesian averaging of binary models which allows us to take into account the uncertainty regarding the set of regressors.

We find that institutional indicators play an important role in all currency crises considered, thus providing persuasive evidence in favour of the "wake up call" hypothesis for financial contagion. Trade, distance, legal system, and financial indicators may also play a role in determining the direction of contagion, but their importance depends on the nature of the crisis and may also be sensitive to the specification of the prior.

We carefully construct a prior density that contains relatively little prior information, while still allowing for simple computations. We also carry out simulations to illustrate that classical tests of significance in probit models often have unacceptable error rates, and that Bayesian methods provide a reliable alternative.

JEL Classification: F31, F32, C11

Keywords: Financial contagion; Speculative attack; Exchange rate; Institutions; Bayesian Model Averaging; Probit Models; Prior Elicitation

1 Introduction

Currency crises tend to occur in waves. In repeated instances from the early 1970s to the late 1990s it has been observed that when speculative attacks lead to a crisis in one country, market volatility tends to spread to other countries in the region and elsewhere. Several authors have examined the possible mechanisms that drive this spread, identifying trade and financial links, changes in risk aversion and "wake-up calls" regarding the sustainability of specific institutions and development models as potential channels by which contagion spreads among countries.

In this paper we empirically evaluate the relative importance of the various potential transmission mechanisms that have been proposed in the existing literature, by analysing four waves of currency crises in the 1990s. We make two contributions. First, we simultaneously include institutional variables alongside the trade, finance and macroeconomic variables commonly analysed in empirical literature on contagious currency crises, thereby directly testing the "wake-up call" hypothesis. Second, we utilize Bayesian methodologies hitherto unused in the empirical literature on contagion to overcome data limitations and model uncertainty. In particular, we use Bayesian averaging of binary models, which allows us to take into account the uncertainty regarding the set of regressors that should be included. This uncertainty inevitably arises from the existence of alternative hypotheses that could potentially explain contagion of currency crises.

At this stage it is also important to note what we do not do in this paper: we do not seek to enter the debate on whether contagion exists.[1] It is impossible to "prove" that waves of currency crises are not caused by some common shock in some unobserved (and possibly unknown) variable or combination of variables. However, it seems unlikely that scattered currency crises can be attributed to global shocks and most market participants agree on the existence of contagion. Micro-theoretic work by Allen and Gale (2000), Morris (1997), Goldstein and Pauzner (2004), and Dasgupta (2004), among others, provide (stylised) structural backing for the existence of contagion in market equilibrium. We therefore *assume* that contagion exists and aim only to shed light on the mechanisms by which it may occur.

Much of the literature on contagious currency crises stresses the phenomenon of regional contagion. It has therefore focused on trade and financial links, which also occur in geographical clusters. Evidence is found in favour of both trade competition and financial links as potential transmission mechanisms for contagion.[2] However, given the high correlation between variables capturing competition for funds and those measuring trade links and trade competition, the results on the importance of the financial variables are not always robust to the inclusion of trade variables. Our paper provides guidance on which of these channels is in fact the most important one.

Moreover, the currency crises of the 1990s have spread far beyond the region of the original crisis country. Glick and Rose (1999) deem that Hong Kong, Indonesia, the Philippines and Thailand were affected by the "Mexican crisis" in 1994 / 1995, while Argentina, Brazil, the Czech Republic, Hungary and South Africa are considered to

[1] See Dungey and Tambakis (2003) for a discussion of the term "contagion" as well as Dungey *et al* (2003) for a detailed review of the contagion literature.

[2] Eichengreen *et* al (1996), Glick and Rose (1999), Kaminsky and Reinhart (2000), Caramazza et al (2000), Van Rijckeghem and Weder (2003).

have been among the victims of the Asian Crisis. According to Van Rijckeghem and Weder (2001) the Russian crisis of 1998 affected 16 countries outside the former Soviet Union, including Argentina, Hong Kong, Indonesia, South Africa and Turkey. While trade competition in third markets may be a possible explanation of extra-regional contagion, it is also interesting to examine the possibility that a speculative attack on a country follows from a "wake-up call" highlighting vulnerabilities associated with particular institutional features, which are found in other countries outside the region.

Bayesian methods are particularly relevant to the analysis of binary data in the context of contagious currency crises for a number of reasons. First, empirical samples are of necessity small: in all previous studies the number of observations is below 100 countries. Of these only a small subset of which experience a crisis in each episode of contagion. Unlike maximum likelihood, Bayesian methods are also valid in small samples. Moreover, Bayesian Model Averaging allows us to obtain results that are robust to the specification of the regression equation. In particular, we appropriately take into account the uncertainty regarding the set of variables that actually contribute to explain the spread of currency crises. We illustrate the performance of Bayesian and Maximum Likelihood tests of significance with a Monte Carlo experiment.

We examine data on currency crises in 1992, 1994, 1997, and 1998, focussing on the relative importance of trade, financial links, institutional features, and macroeconomic variables on the direction of contagion. We report the following results:

1. Institutional variables play an important role in all currency crises considered, thus providing persuasive evidence in favour of the "wake up call" hypothesis for financial contagion. We find that "institutional similarity" in terms of the quality of governance increases the likelihood of contagion in years 1994, 1997 and 1998.
2. Trade, distance, legal system, and financial indicators might also play a role in determining the direction of contagion, but their importance depends on the nature of the crisis. The relevance of these variables is also sensitive to the specification of the prior.

Our paper is linked to a large and growing literature on financial contagion. In what follows, we briefly survey this literature.

2 Literature review

The literature has considered a number of potential channels for international financial contagion.[3] The first potential channel derives from international trade.[4] If a country experiences a sharp devaluation it gains a competitive advantage over its trade partners and over competitors in third markets. To the extent that (the expectation of) deteriorating current account deficits signals potential currency weakness, countries with strong trade connections to the "ground zero country" become more likely to experience a speculative attack. Glick and Rose (1999) examine the importance of the trade channel and find statistical evidence from cross-country data that currency crises spread among countries which have strong trade

[3] See Pericoli and Sbracia (2001) and Dungey *et al* (2003) for literature reviews. Chui *et al* (2004) sets out the framework for assessing external vulnerabilities in more detail.
[4] For a theoretical formalization of this idea see, for example, Gerlach and Smets (1995).

links. However, trade integration is particularly strong within regions, which also tend to have strong financial linkages, as well as cultural and institutional similarities. It is possible that proxies for trade integration also proxy for these other similarities.

A second potential channel of contagion derives from financial linkages between countries.[5] Here contagion arises because groups of countries rely on common creditors and investors. If a country experiences a speculative attack, its major creditor banks may experience liquidity problems, which undermine their ability to provide emergency finance to other countries or trigger capital outflows to restore capital adequacy ratios. Therefore, countries which rely on external funding from the same creditors and investors as the "ground zero country" become vulnerable to speculative attacks. The importance of the "common creditor effect", meaning contagion through bank lending, has been empirically examined by Van Rijckeghem and Weder (2001 and 2003), Caramazza *et al*. (1999), Hernandez and Valdes (2001) and Kaminsky and Reinhart (2000). The results indicate that vulnerability to speculative attacks can spread among clusters of countries which depend on the same lenders. Caramazza *et al*. (1999) additionally show that countries which are more important to the common lenders are more likely to become crises countries than those which only receive a very small proportion of the common lenders' total lending. However, a potential source of concern regarding these results is that the clusters of countries which depend on common lenders – e.g. South America on US banks, South East Asia on Japanese banks and Eastern Europe on German and other European banks – also have either strong trade links or are in direct trade competition with each other. It is therefore difficult to disentangle the relative importance of the trade and the financial channel using classical regression methodologies.

A third channel for contagion derives from shared updating by market participants about the sustainability of specific institutional frameworks or development models. Such a view of contagion is commonly referred to as the "wake-up call" hypothesis.[6] The argument here is that if a country with a particular development strategy, institutional set-up or macroeconomic situation experiences a devaluation, this may be seen as revealing information about the vulnerability of countries of a similar "type" and hence cause the spread of crises.[7] A good example of a major re-evaluation of an economic development strategy was seen in the rapid turn-around in 1997 from applauding the "Asian Miracle"[8] to deploring the "Asian Debacle". Months before the crisis South East Asia's "dedicated capitalism"[9] and "Asian values" were praised and held up as strategies for successful development the world over, but were swiftly condemned as "crony capitalism" in the immediate aftermath of the crisis and held responsible for economic vulnerabilities. Issues such as "corruption", "regulatory quality" and "transparency" suddenly came to the forefront of investor attention and may have contributed to the spreading of the crisis to countries perceived to have

[5] For theoretical models formalizing this hypothesis, see, for example, Goldstein and Pauzner (2005), Allen and Gale (2000), and Dasgupta (2004).
[6] For a theoretical model formalizing this hypothesis, see Rigobon (1998). Van Rijckeghem and Weder (2003) provide evidence for the "wake-up call" hypothesis from the Russian crisis, which caused generalized outflows from emerging markets.
[7] See Drazen (1998) on "information externalities"
[8] See for example the 1993 World Bank publication "*The South East Asian Miracle*" hailing the "fundamentally sound development policies" and "tailored government interventions" in eight high performing Asian economies.
[9] Porter (1996)

similar deficits in accountability and data quality. While a large literature has emerged in recent years to measure and quantify the effects of legal and institutional variables on financial development[10] and financial fragility[11] to our knowledge no direct test of the impact of institutional similarity on financial contagion has been carried out. It is a contribution of this paper to provide a direct examination of the "wake-up call" hypothesis using measures of institutional similarity provided in the literature.

3 Data

In Table 1 we summarize the variables that we use. To examine the relative importance of the trade channel we use the "trade share" indicator computed by Glick and Rose for 1992, 1994 and 1997 and Van Rijckeghem and Weder (2003) for 1998. A high value of this index indicates that the country's exports compete intensely with the ground zero country in third markets.

For financial variables, we choose two indicators of competition for funds based on Caramazza *et al*. (1999). Define the "common lender" to be the creditor country most exposed to the ground zero country. For any given country, our first indicator indexes the importance of the common lender to that country. For the emerging market crises the "common lenders" are the US (1994), Japan (1997) and Germany (1998). For example, in the Russian crisis of 1998 the indicator looks at the proportion of country i's total borrowing which derived from German banks. Our second indicator measures how important a potential target country is to the common lender. Thus, the indicator measures country i's borrowing as a proportion of the total loans made by the common lender. We also include a multiplicative interaction of these two indicators. The data are taken from the Bank for International Settlements' (BIS) consolidated data, covering bank lending from banking systems in the "reporting area" of 18 industrialised countries to countries outside the "reporting area".[12] All indicators refer to banks' position reported at the date closest to the respective crises i.e. December 1994 for the Mexican crisis, June 1997 for the Asian crisis and June 1998 for the Russian crisis. The BIS data only cover lending from the reporting area to countries outside the reporting area, meaning that no financial data are available for the 1992 crisis in the European exchange rate mechanism. However, it is likely that contagion through financial centres is a phenomenon limited to emerging market currency crises.

Our analysis of the "wake-up call" hypothesis is based on a number of variables that have been used in the literature to capture institutional similarity between countries. For our main set measures of institutional similarity, we use a number of variables taken from the set of governance indicators compiled by Kaufman *et al*. (1999) for the World Bank. In particular, we test whether variables such as corruption, regulatory quality, and the degree to which the rule of law is upheld influence whether investors withdraw capital from a country. A disadvantage of this dataset is that data collection

[10] See Beck and Levine (2003) for a review

[11] Demirgüç-Kunt and Detragiache. (1998), Kaminsky and Reinhart (1999), Kaminsky and Schmukler (2003)

[12] The reporting area countries are: Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Luxembourg, Netherlands, Norway, Spain, Sweden, Switzerland, UK and US

only began in 1996. However, Kaufmann *et al.* (2005) note that the quality of governance tends to be highly persistent, institutions change only slowly.[13] Changes in governance over time are small relative to the level of the governance indicators and the reported error margin on the estimates. Changes in annual governance estimates where the 90% confidence intervals do not overlap are only reported in a tiny minority of countries.[14] We therefore take the average score of each country in the years 1996, 1998 and 2000 and used this for each episode of the 1990s currency crises. We then express these variables with respect to the ground zero country. Let c_i be the corruption index for country i that is constructed as just described, and let c_0 be the same for the ground zero country. Then the variable that we use in our analysis is defined as: $|c_i - c_0|/|c_o|$. An analogous index of similarity is constructed for the other two institutional variables.

An additional way of capturing institutional similarity derives from legal origin. The large literature on law and finance (e.g. La Porta *et al*. 1998) argues that a country's legal system (mostly acquired through colonisation or occupation) has important effects on how confidently investors transact in a country, and that this differs significantly between Anglo-Saxon common law and French, German and Scandinavian civil law systems.[15] Motivated by this literature, we complement our core measures of institutional similarity summarized above by an indicator of common legal origin, which takes the value 1 if a country has the same legal system as the ground zero country. The data are taken from La Porta *et al*. (1998).

In addition to these variables we use the relative geographical distance to the ground zero country. Countries that are closer are likely to have more similar institutions and culture. Therefore, in the minds of currency speculators, they might be more attractive potential victims in a wave of currency attacks. The distances between countries were computed as the distances between capital cities, using the distance calculator provided by Darrell Kindred[16] at http://www.indo.com/distance.

Finally, we proxy for the overall level of institutional development using GDP/capita. This tends to be highly correlated with indicators of institutional quality.

We use a number of macro-economic variables as "control variables" in the regression, such as current account and budget deficits, countries' reserve positions, credit expansion, inflation and growth performance. These variables control for the possibility that a country would have fallen into crisis regardless of the attack on the first country, because of its own weak macroeconomic fundamentals.[17] In our choice of control variables, we have been guided by the prior work of Eichengreen *et al.* (1996), Glick and Rose (1999) and Van Rijckeghem and Weder (2003). The variables are computed or taken from the IFS for the period preceding the crisis[18]. This reflects both the delay in data becoming available and the fact that in the

---

[13] http://www.worldbank.org/wbi/governance/pdf/GovMatters_IV_main.pdf
[14] http://www.worldbank.org/wbi/governance/pdf/govmatters3_wber.pdf
[15] See Beck *et al* (2001) for a review
[16] This calculator uses the latitudes and longitudes of the cities concerned and then computes the distance between them by using the Geod program, which is part of the PROJ system, a set of cartographic projection tools, provided by the US Geological Survey at ftp://kai.er.usgs.gov/pub/.
[17] See e.g. Kaminsky *et al* (1998) for a review of the empirical currency crises literature
[18] 1994 for Mexico, 1996 for Asia and 1997 for Russia

immediate aftermath of a currency crisis there usually is significant worsening of the macroeconomic situation.

4 Methodology

4.1 Models

Let Z be the $n \times k$ matrix that contains all variables that could potentially enter in the regression equation, where n is the number of observations and k is the number of potential regressors. Let $Y = (y_1,...,y_n)'$ be an $nx1$ vector of observed binary variables. We consider all binary probit models that result from including a different subset of Z as explanatory variables, and this gives rise to 2^k models. In particular, model M_j is defined as the following probit model:

$$Y^* \mid \theta \sim N\left(Z_j^{'}\theta_j, I_n\right)$$
$$y_i = I_{[y_i^*>0]}$$

where $Y^* = (y_1^*,...,y_n^*)'$ is an $nx1$ vector containing unobserved latent data, Z_j is a $n \times k_j$ submatrix of Z, θ is a $k \times 1$ vector of unknown parameters, θ_j is a $k_j \times 1$ subvector of θ containing the elements of θ that are not restricted to be zero in model M_j, I_n is the identity matrix of dimension n and $I_{[.]}$ is an indicator function which takes value one if the expression in brackets is satisfied and zero otherwise. Note that the elements of θ that are restricted to be zero in model M_j (and are therefore not included in θ_j) will be unrestricted in other models.

Our inference for θ is based on the posterior mean and credible regions[19] of the posterior density of θ ($\pi(\theta \mid Y)$), which is a weighted average of the posterior densities obtained under each of the models ($\pi(\theta \mid Y, M_j)$):

$$\pi(\theta \mid Y) = \sum_{j=1}^{2^k} \pi\left(M_j \mid Y\right)\pi\left(\theta \mid Y, M_j\right)$$

where $\pi\left(M_j \mid Y\right)$ represents the posterior probability of model M_j. To give an intuitive idea of what this posterior probability represents, we note that when the number of observations is sufficiently large, the posterior probability of each model is proportional to the exponential of the Bayesian Information Criterion (Schwarz, 1978), which is equal to the maximised log-likelihood penalised by the number of parameters in the model. For a precise definition of how the posterior probability of a model is defined, see Koop (2003, p.p. 23-26).

[19] A 95% credible interval is the Bayesian analogue of a frequentist 95% confidence interval, and it is an interval that contains the true value of the parameter with probability 95% (e.g. see Koop 2003, p. 44).

Let us define the probability of inclusion for a group of explanatory variables S_j (possibly containing just one variable) as the joint probability of all models that include at least one of the variables in S_j. In other words, the probability of inclusion of S_j is the probability that at least one variable in S_j has a non zero effect on the expected outcome of the dependent variable. Thus, a zero inclusion probability implies that all of the coefficients in θ that correspond to S_j are equal to zero.

The Bayesian methodology presents a number of important advantages over its more commonly used classical counterparts in the context of the contagion literature. Firstly, it allows us to compare simultaneously all possible models. This is in contrast with classical methods, in which only two hypotheses are compared at a time. Secondly, the conditional error rate in significance testing is under control (e.g. Berger and Selke 1987, Selke, Bayarri and Berger, 2001). To give an illustration of this second advantage and test the robustness of Bayesian methods to prior specification in the context of Probit models, we provide a Monte Carlo experiment in the appendix.

4.2 Prior

We propose a prior that is both convenient for performing calculations and that is relatively non-informative. For computational simplicity, we choose a normal prior for the coefficients of model M_j:

$$\theta_j \mid M_j \sim N(0,V), \qquad V_j = g(Z_j^{'}Z_j)^{-1},\ g > 0 \tag{1}$$

A prior mean equal to zero implies that we consider outcomes y_i=1 and y_i=0 to be equally likely a priori for i=1,..,n. This prior mean and prior correlation matrix have been chosen for Bayesian estimation and model selection by many other authors (e.g. Zellner, 1986, Poirier 1985, Fernandez Ley and Steel, 2001). In the following we discuss prior elicitation for g. We will consider three choices for g, and obtain empirical results for each of them.

We first note that simply choosing a very large value for g would not result in a reasonable prior. Such a large value of g implies that a priori we expect the probability of (y_i=1) to be either 1 or zero for every i=1,…,n and consequently marginal effects (on probabilities) to be approximately zero. That is, probabilities between 0 and 1 would receive very small prior weight and marginal effects near zero would receive very high prior weight.[20] Therefore, instead of simply fixing a very large value for g, we adapt for our exercise priors that have been proposed for other related models.

[20] For example, suppose that there is only one regressor in the model and no constant term. A sufficiently large prior variance for the slope coefficient implies that the probability that $Z\theta$ is in the interval (-4,4) is approximately zero. Note that in order to predict the outcome of y_i it does not matter in practice if $Z\theta$ is –5 or –250, since both values result in the probability of y_i=1 being approximately equal to zero. Therefore, since a large prior variance effectively rules out that $Z\theta$ lies in (-4,4), the size of the slope coefficient is no longer relevant, and all we would need, should the prior information be true, in order to predict *perfectly* the outcome of y_i, is the sign of the slope coefficient. Thus, because the prior would be so informative, the only relevant information that we would expect from the data would concern the sign of the slope coefficient. A large amount of data would be necessary to change such strong prior beliefs on large probabilities and small marginal effects.

Our first choice of prior fixes a value of g such that:

$$Var(\bar{z}_j^{'}\theta_j) = \bar{z}_j V_j \bar{z}_j^{'} = 1$$

where $\bar{z}_j$ is a k_j×1 vector containing the average sample values of Z_j. This implies the following value of g:

$$\bar{g} = \left(\bar{z}_j^{'}\left(Z_j^{'}Z_j\right)^{-1}\bar{z}_j\right)^{-1} \qquad (2)$$

To see why this choice of g is appealing, let us define $\pi = \Pr(y=1 \mid \bar{z}_j, \theta_j, M_j) = \Phi\left(\bar{z}_j^{'}\theta_j\right)$, where Φ is the distribution function of a standard normal and therefore π is the probability of (y=1) for a country with average values for the regressors. If we fix g to be equal to $\bar{g}$, then our prior for π is a uniform in the interval (0,1). Geisser (1984) argues that, in the context of estimating a probability with a binomial likelihood, a uniform prior has more compelling characteristics than other form of priors[21].

Another popular choice of non-informative prior to estimate a probability is a Beta distribution with both parameters equal to ½ (i.e. $\pi \sim$ Beta(1/2,1/2), Lee, 1987, p.241). This prior, in the context of a binomial likelihood, is non-informative according to alternative criteria used by different authors (Jeffreys, 1961, Box and Tiao, 1973, Akaike, 1978 and Bernardo, 1979). Compared to the uniform prior, the Beta prior gives slightly more weight to values near to zero and near to 1. In our model, this implies that values of θ_j that were further away from zero would receive greater prior weight. Within our framework, we can achieve this greater weight by choosing $g = a\bar{g}$, with a>1. After experimenting with several values for a, we found that a=2.46 results in a prior for π that approximates well a Beta(1/2,1/2). This is illustrated in Figure 1, which shows that our prior for π when a=2.46 is virtually undistinguishable from the Beta prior. Therefore, the second prior that we consider results from fixing $g = 2.46\bar{g}$. Finally, for sensitivity analysis we will also consider prior (1) with $g = 5\bar{g}$.

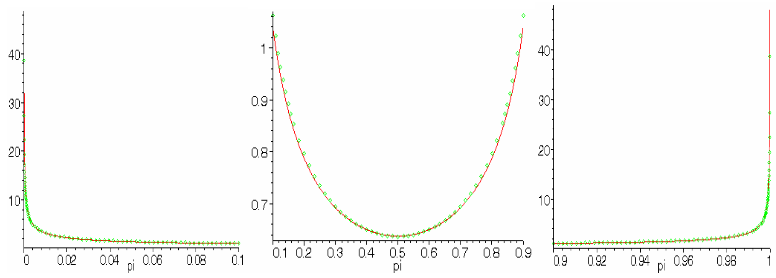

Figure 1: Three views of our prior density for π with $g = 2.46\bar{g}$ (continuous line) and a Beta(1/2,1/2) (dotted line).

[21] In addition, we note that if Z contains an intercept term, then expression (2) is equal to n. A value of g equal to n has been recommended in the context of model selection by Fernandez, Ley and Steel (2001).

4.3 Computation

In order to obtain the posterior probabilities of inclusion, posterior means and credible intervals we use the algorithm proposed by Holmes and Held (2003), who adapt the algorithm of Albert and Chib (1995) to allow for model uncertainty. The algorithm is a Markov Chain that visits a model at each iteration n (M_n), and also at each iteration generates a value for θ conditioning on M_n. Even though the model and value of θ at the initial iteration is chosen arbitrarily, as the number of iterations increases, the models and parameter values generated can be regarded as a sample from the posterior distribution. Therefore, posterior means and other quantities of interest can be easily approximated with their sample analogues. We briefly describe the algorithm here.

Let M_n be the model visited in the n^{th} iteration of the Markov Chain algorithm, let θ_n be the value of the non-zero parameters in M_n at the n^{th} iteration and similarly let Y_n^* be the value of Y^*. Assuming prior (1) for the parameters in a model, and assuming that all possible models have equal prior probabilities, the (n+1) iteration proceeds as follows:

1) Choose a model M^* from a uniform distribution defined on the following set of models:
    - Model M_n
    - Models that result from dropping one regressor in M_n
    - Models that result from adding one regressor to M_n
2) Set M_{n+1} equal to M^* with probability:

$$\alpha = \min\left\{1, \frac{(1+g)^{-k_*/2}\exp\left(-1/2\left(Y_n^*\right)'\left(I_n - 1/(1+g)Z_*V_*Z'_*\right)Y^*\right)}{(1+g)^{-k_n/2}\exp\left(-1/2\left(Y_n^*\right)'\left(I_n - 1/(1+g)Z_nV_nZ'_n\right)Y^*\right)}\right\}$$

   where I_n is the identity matrix of dimension n, Z_* is a $n\times k_*$ matrix with the set of regressors contained in M^*, k_n is the number of regressors in M_n and V_* and V_n are defined as in (1). Set M_{n+1} equal to M_n with probability 1-α.
3) Draw each of the components of Y_{n+1}^* from univariate truncated normal distributions as explained in Albert and Chib (1995).
4) Draw θ_{n+1} from a normal density with covariance matrix ($\tilde{V}$) and mean ($\tilde{\mu}$) equal to:

$$\tilde{V} = \frac{g}{g+1}\left(Z'_{n+1}Z_{n+1}\right)^{-1} \qquad \tilde{\mu} = \tilde{V}Z'_{n+1}Y_{n+1}^*$$

   where Z_{n+1} is the set of regressors that are included in model M_{n+1}.

We calculate the posterior probability of model M_j as the proportion of iterations that visit model M_j. Similarly, posterior means and credible intervals for θ or functions of θ (e.g. marginal effects) can be calculated using the draws obtained with the algorithm.

5 Results

In Table 1 we list the definitions of the variables that we use. Our main economic results are presented in Tables 2 and 3. Tables 4 and 5 assess the out-of-sample predictive ability of the models. The dependent variable is binary, taking value one if the country concerned suffered a crisis. For each variable we report three quantities. First, we report the probability of inclusion of the variable (p), as defined in Section 4.1. This is the probability that the effect associated with a regressor is different from zero. Second, since Probit coefficients are hard to interpret, we report the posterior mean for the marginal effect of each variable. These marginal effects are evaluated at the sample mean of variables.[22] Third, for each marginal effect, we include the 95% credible interval, as defined in Section 4.1. This is the Bayesian analogue to the classical 95% confidence interval in a Maximum Likelihood estimation. Finally, we report the *joint* inclusion probability for the institutional similarity variables (*R. Law*, *Reg. Q.* and *Corrupt*) and for the finance variables (*Fi*1, *Fi*2 and *Fi*1**Fi*2). Since our goal is to understand whether trade competition, financial links, or institutional similarity drive financial contagion, it is important for us to compare the joint probabilities of inclusion of these different categories of variables.

The results reported in Tables 2 and 3 correspond to the prior with $g = \bar{g}$ 2.46. The results that we comment upon are robust to the 3 choices of g, unless otherwise stated.[23]

Institutions

The main conclusion from our empirical analysis is that institutional similarity is an important predictor of financial contagion during emerging market crises. With our two core priors, the joint probability of inclusion of the institutional similarity variables is at least 95% in all crises episodes with the exception of 1992. When the prior has $g = 5\bar{g}$, the joint probability of institutions in 1998 is still high but decreases to 90%. In 1992 the joint probability is above 80%, which is high but not conclusive. Credible intervals at 95% for the marginal effects of institutional variables almost always exclude positive values, which is consistent with the wake-up call theory: countries that are institutionally similar to the ground zero country are more likely to experience crises. The only exceptions are 95% credible intervals for *R. Law* and *Reg. Q.* in 1998, which contain positive values. However, the effects of these two variables in 1998 are more likely to be zero, since their inclusion probabilities are only 27% and 22%, respectively.

Since it is difficult to interpret the size of the marginal effects of the institutional similarity variables, we now provide an alternative way of assessing whether the estimated effects are large or small. Consider a country A that has average value for all regressors except for the institutional similarity variables (*R. Law*, *Reg. Q.* and *Corrupt*), all of which take value 0: that is, this is a country that is very similar to the ground zero country with respect to institutions. In addition, consider a country B that also has average value for all regressors, but whose institutional variables take the same value as the country in our sample that is most dissimilar, in terms of

[22] Note that since we have a dummy variable among the regressors, namely Legal Origin, by taking the sample mean of variables we are evaluating the marginal effect at the average intercept. The marginal effect for the dummy variable Legal Origin is calculated as the change in probability when Legal Origin changes from 1 to 0. The marginal effects for the finance variables (Fi1 and Fi2) take into account the consequent change in the interaction variable Fi1*Fi2.

[23] Results with the other 2 priors are available from the authors upon request.

institutions, to the ground zero country[24]. Hence, countries A and B are different only with respect to institutions. Then, country A is affected by the crisis in years 1994, 1997, 1998 with probabilities (28%, 63%, 55%), whereas the corresponding probabilities for country B are zero for each year.[25] This confirms that institutional similarity played a particularly important role in the direction of spread of the emerging market crises of 1994, 1997 and 1998.

Our results on the effects of common legal origin are less emphatic. Zero values can never be confidently ruled out for the effect of *Legal Or* in any of the crises, especially in 1992, in which the effect of this variable seems to be negligible. The probability of inclusion of *Legal Or.* is highest in year 1997, in which positive values can be ruled out, indicating that countries with the same legal system as the ground zero country experienced lower probability of crisis. The 1997 ground zero country has British legal origin, which suggests that overall countries with British legal origin were less susceptible to financial crises, which is consistent with the results of the Law and Finance literature.[26] The opposite effect is observed in years 1994 and 1998, where the ground zero countries have French and (Post-Socialist) civil law legal origins respectively. However, in these years the probability of the effect being zero is even higher.

Finally, the GDP/capita variable, which captures the idea of groups of countries at the same stage of development being attacked, has a low probability of inclusion in all the regressions.

We now turn to the other potential channels for financial contagion. Our results suggest that after controlling for institutional similarity, other variables such as financial linkage, trade competition and distance have limited impact. We provide a detailed discussion in what follows.

Finance

The joint probability of inclusion of finance variables is above 90% only for the 1998 crisis, provided that the prior variance g is equal to $\bar{g}$ or $2.46\,\bar{g}$. This probability decreases to 84% when $g = 5\,\bar{g}$. Despite the high joint probability, the individual inclusion probabilities of $Fi1$, $Fi2$ and $Fi1{*}Fi2$ are low. This is probably caused by multicollinearity. Despite of the problem of multicollinearity, it can be observed that the marginal effects of $Fi1$ and $Fi2$ in 1998 are positive, since credible intervals exclude negative values. Furthermore, the size of the mean marginal effects is non-negligible. Although the effect is not as clear for other years, the evidence for 1998 confirms the intuition that the more dependent the country is on the common lender, the more likely that it will be affected by the crisis.

---

[24] The most dissimilar country in our sample is defined as the country that maximises the Euclidean distance with respect to the ground zero country. In terms of the variables that are defined in Table 1, it maximises $(R.\ Law)^2 + (Reg.\ Q)^2 + (Corrupt)^2$. According to our data, the most dissimilar countries to the ground zero countries (in terms of institutions) for 1992 (Finland), 1994 (Mexico), 1997 (Thailand) and 1998 (Russia) were Guinea-Bissau, Singapore, New Zealand and Singapore, respectively.

[25] For 1992 this probability decreases from 8% to zero.

[26] See Beck *et al* (2001)

Trade and Distance

The inclusion probability of *Trade* is highest in the 1997 crisis. It is 94% when $g = 5\bar{g}$, but it is below 90% when g is equal to $\bar{g}$ or $2.46\bar{g}$ (and it then takes values 81% and 86%, respectively). However, 95% credible intervals indicate that the possibility of negative values can be confidently neglected, and that mean marginal effects are sizeable, indicating therefore that the trade channel of contagion was probably important in 1997. Although zero values are even more likely in 1994 and 1998, 95% credible intervals indicate that moderately large effects are still possible, and negative values are very unlikely. Therefore, *Trade* could also have been an important determinant in years 1994 and 1998. However, in contrast with Glick and Rose (1999), we find that *Trade* has a negligible effect in the 1992 crisis. It is *Distance*, instead, that seems to play an important role. *Distance* in 1992 is probably simply capturing the fact that EMU countries, which happen to be geographically near, were much more likely to be affected by the crisis. However, the negligible effect of trade is not caused by accounting for distance: if we excluded distance from the set of *potential* regressors the marginal effect of trade would continue to be small.

Out of Sample Predictions

We evaluate the predictive performance when we use the prior with $g=\bar{g}\,2.46$ and the following predictive rule, which is defined for several values of p:

- y_i is predicted to be one when the posterior mean of $\Pr(y_i = 1|Z)$ is greater than p.
- y_i is predicted to be zero when the posterior mean of $\Pr(y_i = 1|Z)$ is smaller than p.

Predictions are made for (1997, 1998) based on parameter estimates from 1994 data. Similarly, predictions are made for (1994, 1998) based on parameters estimated with 1997 data, and for (1994, 1997) based on 1998 data. For each of these three cases we calculate two error rates: E_0 is the proportion of observations that were predicted to be zero but were actually 1. Similarly, E_1 is the proportion of observations that were predicted to be one but were actually 0. Tables 4 and 5 show the results.

Table 5 shows that E_0 is equal to 0 when p is 0.05, and it is smaller than 0.1 when p=0.10. This suggests that the model produces reliable predictions of zeros, in the sense that a small posterior mean of $\Pr(y_i = 1|Z)$ could be taken as strong evidence against the occurrence of a crisis.

Table 4 shows that there are few countries for which the posterior probability of a crisis is high, and this introduces a small sample bias in our estimate of E_1. For example, if 1994 data was used to predict 1997-98 crises, only four cases would have a posterior probability of a crisis greater than 0.90, all of which did actually suffer a crisis: Indonesia in 1997, Republic of Korea in 1997, Malaysia in 1997 and Brazil in 1998. Therefore, the error rate E_1 equals 0, thereby suggesting that predictions of ones are also reliable. However, given the small number of cases that are predicted to be 1, the estimate of this error rate is bound to be imprecise. This is confirmed by the fact that calibration with 1997 or 1998 data yields larger prediction error rates.

6 Conclusions

We have evaluated several hypotheses on the spread of currency crises. We looked at trade, financial linkages, legal and institutional indicators, geographical distance and included other macroeconomic variables as controls. We used Bayesian model averaging of probit models to obtain estimates and significance tests that are robust to the set of variables that is included in the regression equation. Uncertainty regarding the set of regressors in the regression equation is an inevitable consequence of the existence of numerous hypotheses on the spread of currency crises.

We carefully elicited values for the prior variance that resulted in relatively little information in the prior, while still allowing for simple computations. Simple simulations were also carried out to illustrate that classical tests of significance in probit models often have unacceptable error rates, and that Bayesian methods provide a reliable alternative.

We found that institutional variables played an important role in all currency crises considered, and that similarity to the ground zero country in terms of perceived corruption levels increases the likelihood of contagion in years 1994, 1997 and 1998.

Results suggested that trade, distance, legal system and financial indicators might also play a key role on contagion, but their importance depends on the nature of the crisis and sometimes on the specification of the prior.

Bibliography

Akaike, H, (1978), "A new look at the Bayes Procedure," *Biometrika*, 65, 53-59.

Albert, J.H. and Chib, S. (1995), "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association* 88, 669-679.

Allen, F. and Gale, D. (2000), "Financial Contagion," *Journal of Political Economy*, 108, 1-33

Beck, T., and Levine, R.E. (2005), "Legal Institutions and Financial Development" in C. Menard and M. Shirley, (eds.), *Handbook of New Institutional Economics*, Kluwer Dordrecht (The Netherlands).

Beck, T., Demirgüç-Kunt, A. and Levine, R.E. (2001), "Legal Theories of Financial Development" *Oxford Review of Economic Policy*, 17, 483-501

Berger, J. and Selke, T. (1987), "Testing of a point null hypothesis: the irreconcilability of significance levels and evidence (with discussion)," *Journal of the American Statistical Association*, 82, 112-139.

Bernardo, J.M. (1979), "Reference Posterior Distributions for Bayesian Inference" (with discussion), *Journal of the Royal Statistical Society*, B, 41, 113-147.

Berndt, E.R. and Savin, N.E., (1977), "Conflict among criteria for testing hypotheses in the multivariate linear-regression model," *Econometrica*, 45, 1263-1277.

Caramazza, F., Luca, A. and Salgado, R. (2000), "Trade and Financial Contagion in Financial Crises," *IMF Working Paper*, WP/00/55.

Dasgupta, A. (2004), "Financial Contagion through Capital Connections: A Model of the Origin and Spread of Bank Panics," *Journal of the European Economic Association*, 2, 1049-1084.

Demirgüç-Kunt, A. and Detragiache, E. (1998), "Financial Liberalisation and Financial Fragility" in *Annual World Bank Conference on Development Economics 1998*, B. Plezkovic and J Stiglitz, (eds.), 303-31.

Drazen, A. (1998), "Political Contagion in Currency Crises," mimeo, University of Maryland.

Dungey, M, Fry, R, González-Hermosillo, B. and Martin, V.L. (2003), "Empirical Modelling of Contagion: A Review of Methodologies," *CERF working paper*, # 8, available at: http://rspas.anu.edu.au/economics/staff/dungey/pdfs/reviewdfgm.pdf

Dungey, M. and Tambakis, D. (2003), "International Financial Contagion: What do we know?" *CERF Working Paper* # 9.

Eichengreen, B, Rose, A. and Wyplosz, C. (1996), "Contagious Currency Crises," *NBER Working Paper* #5681.

Fernández, C, Ley, E. and Steel, M. (2001) "Benchmark priors for Bayesian model averaging," *Journal of Econometrics*, 100, 381-427.

Flood, R.P. and Garber, P.M., (1984): Gold Monetization and Gold Discipline, *Journal of Political Economy*, 92, 90-107.

Geisser, S. (1984) "On prior distributions for binary trials (with discussion)" *American Statistician*, 38, 244-251.

Gerlach, S. and Smets, F. (1995): Contagious Speculative Attacks; *European Journal of Political Economy*, 11, 5-63.

Glick, R. and Rose, A. (1999), "Contagion and Trade: Why Are Currency Crises Regional?" *Journal of International Money and Finance*, 18, 603-617

Goldstein, I. and Pauzner, A. (2004), "Contagion of Self-Fulfilling Curency Crises Due to the Diversification of Investment Portfolios," *Journal of Economic Theory*, 119, 159-183.

Greene, W.H. (2003), *Econometric Analysis*. 5th Ed., Prentice Hall (New Jersey).

Hernandez, L. and Valdes, R. (2001), "What Drives Contagion: Trade, Neighbourhood or Financial Links?" *IMF Working Paper* WP 01/29.

Holmes, C.C. and Held, L. (2005), "Bayesian Auxiliary Variable Models for Binary and Multinomial Regression," forthcoming in *Bayesian Analysis*.

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

Jeffreys, J. (1961), *Theory of Probability*. Oxford University Press (Oxford).

Koop, G. (2003) *Bayesian Econometrics*. Wiley (London).

Kaminsky, G., Lizondo, S. and Reinhart, C. (1998), "Leading Indicators of Currency Crises" *IMF Staff papers*, 45, 1-48.

Kaminsky, G. and Reinhart, C. (1999), "The Twin Crises: the Causes of Banking and Balance of Payments Problems," *American Economic Review* 89, 473-500.

Kaminsky, G. and Reinhart, C. (2000), "On Crises, Contagion and Confusion," *Journal of International Economics*, 145-168.

Kaminsky, G. and Schmukler, S. (2003), "Short-Run Pain, Long Run Gain: The Effects of Financial Liberalisation" *IMF Working Paper* 2003 / 34.

Kaufmann, D., Kraay, A. and Zoido-Lobaton, P. (1999), "Governance Matters," *World Bank Working Paper* # 2196

Krugman, P. (1979) "A model of balance of payments crises" *Journal of Money, Credit and Banking*, 11, 311-325.

Krugman, P. (1998), "What Happened to Asia?" unpublished manuscript.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R. (1998), "Law and Finance," *Journal of Political Economy*, 106, 1113 –1155.

Lee, P. (1987), *Bayesian Statistics: An Introduction*, 1st ed., Oxford University Press (New York).

Morris, S. (1997), "Contagion," *Review of Economic Studies*, 67, 57-78.

Morris. S. and Shin. H.S. (1998), "Unique Equilibrium in a Model of Self-fulfilling Currency Crises," *American Economic Review*, 88, 587-597.

Obstfeld, M. (1994) "The logic of currency crises," *Cahiers Économiques et Monétaires de la Banque de France*, 43, 189-214.

Obstfeld, M. (1996) "Models of currency crises with self-fulfilling features," *European Economic Review* 40, 1037-1047.

Porter, M.E. (1996), "Capital Choice: National Systems of Investment," in *As If the Future Mattered: Translating Social and Economic Theory into Human Behavior*, N.A. Goodwin (ed.), Ann Arbor: University of Michigan Press, 15-44.

Rigobon, R. (1998), "Informational Speculative Attacks: Good News is No News," *MIT Working Paper*

Selke, T., Bayarri, M.J. and Berger, J. (2001), "Calibration of P-values for testing precise null hypotheses," *The American Statistician* 55, 62-71.

Van Rijckeghem, C. and Weder, B.S. (2001). "Sources of Contagion; Finance or Trade?" *Journal of International Economics,* 45, 293-308.

Van Rijckeghem, C. and Weder, B.S. (2003). "Spillovers through banking centers: a panel data analysis of bank flows," *Journal of International Money and Finance*, 22, 483-509.

Trade	Trade competitiveness as defined in Glick and Rose (1999)
Dom. Cred.	Growth of Domestic Credit
Bud/GDP	Budget Position as a percentage of GDP
CA/GDP	Current account position as a percentage of GDP
Growth	Real rate of GDP per capita growth
M2/Res	Ratio of M2 to central bank foreign reserves
Inflation	Domestic CPI inflation
In. GDP	GDP per capita at the beginning of the year measured in 1990 US $ in levels (1992, 1997, 1998) and in natural log (1994).
Distance	Distance to ground zero country in miles (1992, 1997 and 1998) and thousand of miles (1994).
Legal Or.	Legal Origin Dummy: 1 if a country has the same legal system as the ground zero country
R. Law	Similarity, to ground zero country, in the degree to which the rule of law is upheld. Decreasing with similarity. Original data from Kaufmann *et al.* (1999).
Reg. Q.	Similarity, to ground zero country, in Regulatory quality. Decreasing with similarity. Original data from Kaufmann *et al*. (1999).
Corrupt	Similarity, to ground zero country, in Levels of Corruption. Decreasing with similarity. Original data from Kaufmann *et al*. (1999).
*Fi*1	The proportion of a country's total borrowing that was borrowed from the common lender.
*Fi*2	A country's borrowing as a proportion of the total loans made by the common lender.
*Fi*1**Fi*2	The product of Fi1 times Fi2.

Table 1: Definition of variables.

	Crises in 1992				Crises in 1994			
	p	Mean and 95% credible interval for marginal effects			*P*	Mean and 95% credible interval for marginal effects		
		Lower limit	Mean	Upper limit		Lower limit	Mean	Upper limit
Trade	0.29	0	0.011	0.083	0.63	0	0.091	0.482
Dom. Cred.	0.36	0	5.9E-05	3.9E-04	0.25	-1.3E-04	-1.0E-05	8.2E-05
Bud/GDP	0.39	-0.007	-0.001	0	0.52	0	0.001	0.007
CA/GDP	0.33	-1.4E-04	3.1E-04	0.002	0.21	-2.1E-04	8.2E-05	5.1E-04
Growth	0.26	-0.004	-4.4E-04	0	0.21	-1.2E-04	1.1E-04	0.001
M2/Res	0.25	0	8.7E-05	7.4E-04	0.37	-5.3E-04	-7.9E-05	5.3E-06
Inflation	0.43	-0.002	-3.0E-04	0	0.24	-1.7E-05	3.6E-05	2.5E-04
In. GDP	0.23	-7.1E-07	-5.9E-08	3.1E-07	0.42	-0.012	-4.2E-04	0.007
Distance	**0.93**	-3.9E-05	-1.1E-05	0	0.26	-1.2E-04	2.6E-04	0.002
Legal Or.	0.13	-0.002	-5.4E-04	2.0E-05	0.60	0	0.014	0.167
R. Law	0.26	-0.056	-0.003	0.027	0.58	-0.049	-0.009	0
Reg. Q.	0.61	-0.202	-0.041	0	0.34	-0.037	-0.005	0
Corrupt	0.25	-0.050	-0.003	0.027	0.48	-0.038	-0.006	0
*Fi*1	-	-	-	-	0.20	-0.041	-0.002	0.035
*Fi*2	-	-	-	-	0.56	0	0.811	3.941
*Fi*1**Fi*2	-	-	-	-	0.29	-	-	-
Constant	0.42	-	-	-	0.50	-	-	-
	P(*Finance*) = - P(*Institutions*) = 0.82				P(*Finance*) = 0.79 P(*Institutions*) = **0.97**			

Table 2: Probabilities of inclusion, posterior mean and credible intervals for the crises in 1992 and 1994. *p* is the probability of inclusion of each variable. P(*Finance*) is the joint probability of inclusion of *Fi*1, *Fi*2 and *Fi*1**Fi*2. P(*Institutions*) is the joint probability of inclusion of *R. Law*, *Reg. Q.* and *Corrupt*. Prior with $g=\overline{g}$ 2.46

	Crises in 1997				Crises in 1998			
	p	Mean and 95% credible interval for marginal effects			*P*	Mean and 95% credible interval for marginal effects		
		Lower limit	Pos. Mean	Upper limit		Lower limit	Pos. Mean	Upper limit
Trade	0.864	0	0.480	1.257	0.263	-0.098	0.115	0.967
Dom. Cred.	0.392	-0.006	-1.1E-03	0	0.130	-0.004	-1.7E-04	0.001
Bud/GDP	0.620	0	0.012	0.045	0.205	0	0.003	0.030
CA/GDP	0.380	-0.013	-0.002	0	0.178	0	0.002	0.018
Growth	0.353	-0.019	-2.2E-03	0.003	0.187	0	0.005	0.039
M2/Res	**0.952**	6.3E-06	0.011	3.1E-02	0.261	0	0.005	0.034
Inflation	0.225	-1.3E-03	-1.4E-05	0.001	0.152	-0.002	2.8E-04	0.005
In. GDP	0.211	-1.0E-05	-1.1E-06	3.9E-07	0.204	-2.7E-06	3.3E-06	3.5E-05
Distance	0.252	-1.0E-05	-9.4E-07	2.0E-06	0.284	0	8.0E-06	5.3E-05
Legal Or.	0.750	-0.326	-0.136	0	0.367	0	0.175	0.865
R. Law	0.816	-0.233	-0.080	0	0.275	-0.229	0.015	0.343
Reg. Q.	0.428	-0.284	-0.054	0	0.218	-0.040	0.020	0.223
Corrupt	0.217	-0.023	-0.002	9.9E-04	0.807	-0.666	-0.286	0
Fi1	0.218	-0.185	0.099	1.134	0.157	0	0.429	1.253
Fi2	0.196	-0.926	0.103	1.299	0.373	0	6.522	12.378
*Fi1*Fi2*	0.202	-	-	-	0.577	-	-	-
Constant	0.430	-	-	-	0.468	-	-	-
	P(*Finance*) = 0.50 P(*Institutions*) = **0.97**				P(*Finance*) = **0.91** P(*Institutions*) = **0.95**			

Table 3: Probabilities of inclusion, posterior mean and credible intervals for the crises in 1997 and 1998. *p* is the probability of inclusion of each variable. P(*Finance*) is the joint probability of inclusion of *Fi*1, *Fi*2 and *Fi*1**Fi*2. P(*Institutions*) is the joint probability of inclusion of *R. Law*, *Reg. Q.* and *Corrupt*. Prior with $g = \bar{g}$ 2.46

	p	0.5	0.65	0.75	0.85	0.9
1994	E_1	0.58	0.52	0.27	0.17	0.00
	NP	43	29	11	6	4
	AN	27	27	27	27	27
1997	E_1	0.71	0.70	0.70	0.63	0.67
	NP	45	33	20	8	6
	AN	23	23	23	23	23
1998	E_1	0.56	0.29	0.20	1.00	n.a.
	NP	16	7	5	1	0
	AN	18	18	18	18	18

Table 4: Out of Sample Predictions of 1. y_i is predicted to be one when the posterior mean of $\Pr(y_i = 1|Z) > p$. When the models are estimated with 1994 data, predictions are made for (1997, 1998). Similarly, predictions are made for (1994, 1998) based on 1997 data, and for (1994, 1997) based on 1998 data. *NP* is the number of observations predicted to be 1. E_1 is the proportion of *NP* that was actually 0. *AN* is the actual number of ones in the validation sample.

	p	0.5	0.35	0.25	0.1	0.05
1994	E_0	0.14	0.08	0.09	0.04	0
	NP	64	48	43	24	13
	AN	80	80	80	80	80
1997	E_0	0.15	0.13	0.12	0.06	0
	NP	67	60	52	32	16
	AN	89	89	89	89	89
1998	E_0	0.12	0.08	0.05	0.09	0
	NP	89	77	56	34	11
	AN	87	87	87	87	87

Table 5: Out of Sample Predictions of 0. y_i is predicted to be zero when the posterior mean of $\Pr(y_i = 1|Z) < p$. When the models are estimated with 1994 data, predictions are made for (1997, 1998). Similarly, predictions are made for (1994, 1998) based on 1997 data, and for (1994, 1997) based on 1998 data. *NP* is the number of observations predicted to be 0. E_0 is the proportion of *NP* that was actually 1. *AN* is the actual number of zeros in the validation sample.

Appendix: Posterior model probabilities versus p-values in Probit models: an illustration with simulated data

In the context of testing whether some parameters in θ are equal to zero, Bayesian methods offer two important advantages over maximum likelihood (ML) methods. Firstly, Bayesian methods allow for the simultaneous evaluation of numerous hypotheses, whereas ML methods typically compare just two hypotheses at a time. That is, in the language of our previous sections, Bayesian methods allow for the simultaneous comparison of more than 2 models. A second but not less important advantage is that in Bayesian testing, unlike in ML testing, the conditional error rate (*CER*) is under control (e.g. Berger and Selke, 1987).

In order to define the *CER*, assume for simplicity that we consider just two models, or equivalently that we want to test H_0:θ_1=0 versus H_1:$\theta_1 \neq 0$. Assume that we apply repeatedly a testing procedure to different samples that come from *different* populations[27]. Some of these samples come from a population with θ_1=0 and some others from a population with $\theta_1 \neq 0$. When $\theta_1 \neq 0$, assume that the values of θ_1 come from a distribution $f(\theta_1)$. Let p denote the p-value for a particular sample, if the test is frequentist, or let it represent the posterior probability of H_0 in a Bayesian test. Assume that after the test was applied to each sample, it became known the population from which each sample actually came from. Let $R(p)$ be the set of samples that resulted, approximately, in a given value for p. The $CER(p)$, which is defined for each value of p, is equal to the proportion of samples in $R(p)$ that actually came from a population with θ_1=0.

To see why we call this an error rate, consider the case of p=0.05, which is commonly used as a threshold value in hypothesis testing. In this case, R(0.05) will consist entirely of samples where the null hypothesis was rejected (at a 5% significance level in the frequentist case), and *CER*(0.05) captures how many of those samples (as a proportion) were generated from a population where H_0 was actually true. In Bayesian testing, $CER(p)$ is by definition equal to p (i.e. the posterior probability of H_0) when the prior density coincides with $f(\theta_1)$ (Berger and Selke, 1987). In contrast, however, $CER(p)$ can be substantially larger than p in frequentist tests. Indeed, Berger and Selke (1987) show that in the context of testing the significance of the constant term in a linear model with no other regressors but the constant, the *CER*(*0.05*) resulting from a classical test is *at least* 23% and conclude that a p-value of 0.05 should never be considered as enough evidence to reject the null hypothesis. The main message of this and other related work is that 'p-values are commonly thought to imply

[27] Recall that ML procedures are designed to control Type I and Type II error rates, which are defined in the context of repeated sampling from the same population. Type II error rates depend on the true value of θ_1 in the population and the Type I error rate only applies when θ_1=0.

considerably greater evidence against H_0 than is actually warranted' (Selke, Bayarri and Berger, 2001).

As we mentioned, the desired property that the posterior probability of H_0 is equal to *CER*(p) hinges upon the prior being equal to $f(\theta_1)$. We now pass to use simulated data to get a sense of how robust is Bayesian testing to the specification of the prior in the context of Probit models. We also illustrate some of the problems that arise when classical p-values are used as a measure of evidence. Following Selke, Bayarri and Berger (2001), we define some concepts that are related to *CER*(p) but that are workable in a simulation context. We define $R(p_0{:}p_1)$ as the set of samples that result in the posterior probability of H_0 (or p-value in a frequentist test) lying in the interval (p_0,p_1), and define $CER(p_0{:}p_1)$ as the proportion of those samples that came from a population where H_0 was actually true.

Our simulation setup is similar to that in Selke, Bayarri and Berger (2001). We simulate 5000 artificial samples from the following Probit model.

$$y_i^* = z_0\theta_0 + z_1\theta_1 + \varepsilon_i$$

where $\varepsilon_i \sim N(0,1)$ and z_0 and z_1 are two variables that we used in our data analysis, namely trade and inflation in 1992, respectively (see the section on Data for details). We simulate 2500 samples for model M_0, which imposes the restriction $\theta_1=0$, and also 2500 samples for model M_1, in which $\theta_1 \neq 0$. We assume $f(\theta)$, which is the distribution from which the values of θ_0 and the non-zero values of θ_1 are generated, to be equal to (1), with $Z_0=(z_0)$ and $Z_1=(z_0 \;\; z_1)$ and the following 5 values for g:

$$g_1^a = 0.01\overline{g}, \qquad g_2^a = 0.05\overline{g}, \qquad g_3^a = \overline{g}, \qquad g_4^a = 2.46\overline{g}, \qquad g_5^a = 5\overline{g}$$

Therefore, we specify $f(\theta)$ to have the same form as the prior, but look at the effect of mis-specifying the value of g. In particular, we denote the value of g in the prior as g^p and consider, as in the data analysis, three possible values: $g^p_1=\overline{g}$, $g^p_2=2.46\overline{g}$ and $g^p_3=5\overline{g}$. We simulate 5000 samples for each value of g^a in the manner just described. Each sample is then analysed with the three alternative priors: $g^p_1=\overline{g}$, $g^p_2=2.46\overline{g}$ and $g^p_3=5\overline{g}$. Apart from a Bayesian posterior probability, we calculate p-values from likelihood ratio and Wald tests[28] (Greene 2003, section 17.5).

Table 6 shows the values of *CER*(0.01:0.05) for various combinations of (g^p, g^a). In the Bayesian case, when $g^p = g^a$ $CER(p_0{:}p_1)$ should be equal to the average of p in the samples in $R(p_0{:}p_1)$. Therefore, $CER(p_0{:}p_1)$ should lie in the interval $(p_0{:}p_1)$ when $g^p = g^a$. Table 6 shows that Bayesian error rates seem to be quite robust to misspecification of the prior. For instance, if the prior fixes g^p=5, conditional error rates are small even when the non-zero

[28] Whenever Maximum Likelihood estimates did not exist, because regressors predicted the outcome perfectly, we assumed that the tests failed to reject the null hypothesis.

values of θ_1 are generated from a distribution with g^a=1. In addition, if non-zero values are generated with g^a=0.05 then Bayesian conditional error rates continue to be small. However, there is a limit to that robustness. The Bayesian test fails if the value of g^a is 0.01, in which case error rates are in the range of 25%-35%. Note that a very small value of g^a implies that non-zero values of θ_1 are very near to 0, and it is therefore more difficult to detect that they are non zero. With respect to frequentist tests, note that the LR test always have unacceptable error rates for all values of g^a. In particular, the error rate is as high as 43% when g^a=5, and 38% when g^a=2.46. Even though the Wald test performs better, it also has error rates that are always above 10%. Given that p-values are in the range 0.01-0.05, an error rate larger than 10% comes as a surprise. The difference in the performance between these two frequentist tests is probably explained by the different sizes. Even though the nominal size for both tests is 5%, the LR test is slightly oversized (6.6%) and the Wald test is undersized (3.9%)[29]. Note that the size distortion is to be expected, as these tests use asymptotic approximations which are not exact for finite samples (Berndt and Savin, 1977).

[29] The actual sizes were calculated using all available samples: 2500x5= 12500

	g^a=0.01	g^a=0.05	g^a=1	g^a=2.46	g^a=5
g^p=1	0.33 (21)	0.11 (105)	0.024 (247)	0.026 (193)	0.028 (144)
g^p=2.46	0.32 (19)	0.051 (79)	0.034 (207)	0.030 (169)	0.00 (147)
g^p=5	0.25 (24)	0.042 (71)	0.015 (202)	0.023 (171)	0.016 (127)
ML LR	0.44 (324)	0.27 (470)	0.32 (371)	0.38 (305)	0.43 (265)
ML Wald	0.43 (269)	0.23 (466)	0.16 (496)	0.15 (481)	0.11 (479)

Table 6: Summaries of error rates: *CER*(0.01:0.05). The first three rows correspond to a Bayesian test with three different prior variances. The last two rows correspond to LR and Wald tests. The number in brackets is the number of samples in *R*(0.01:0.05), which determines the precision of the estimate of *CER*(0.01:0.05).