WP/04/39

IMF Working Paper

Autocorrelation-Corrected Standard Errors in Panel Probits: An Application to Currency Crisis Prediction

Andrew Berg and Rebecca N. Coke

INTERNATIONAL MONETARY FUND

IMF Working Paper

Research Department

Autocorrelation-Corrected Standard Errors in Panel Probits: An Application to Currency Crisis Prediction

Prepared by Andrew Berg and Rebecca N. Coke[1]

Authorized for distribution by Jonathan D. Ostry

March 2004

Abstract

Many estimates of early-warning-system (EWS) models of currency crisis have reported incorrect standard errors because of serial correlation in the context of panel probit regressions. This paper documents the magnitude of the problem, proposes and tests a solution, and applies it to previously published EWS estimates. We find that (1) the uncorrected probit estimates substantially underestimate the true standard errors, by up to a factor of four; (2) a heteroskedasicity- and autocorrelation-corrected (HAC) procedure produces accurate estimates; and (3) most variables from the original models remain significant, though substantially less so than had been previously thought.

JEL Classification Numbers: C15, C22, C23, C25, F31, F47

Keywords: Currency crisis, early-warning systems, serial correlation, panel probit

Author's E-Mail Address: aberg@imf.org

[1] Portions of this paper were originally written when Rebecca Coke was a Research Assistant in the IMF's Research Department in 1999. Programs are available from the authors on request. We would like to thank Sam Ouliaris for his useful comments.

Contents Page

I. Introduction

Currency crisis models use macroeconomic variables to predict the probability of a currency crisis. Two features of useful currency crisis forecasting models combine to create econometric problems. First, they typically forecast the probability of crisis several months or even years ahead. Such long-horizon forecasts are useful to the extent that they reflect the lags between analysis, policy change, and outcomes. The second feature of a useful system is that it makes use of the highest-frequency data available in updating forecasts. For example, the IMF has attempted to predict the probability of a crisis over a two-year period while updating forecasts monthly.[2]

This combination of features—that is, a forecast horizon longer than the frequency at which the forecast is updated—results in serially correlated prediction errors in situations in which both the dependent variable and the explanatory variables are themselves serially correlated.

Serially correlated errors can present a variety of problems. In the models we have in mind, the dependent variable is defined as a {0 1} outcome, and the observations are from a panel of countries. For nonlinear models such as a probit, coefficient estimates are consistent, as long as the model is otherwise correctly specified.[3] The standard errors, however, will, in general, be biased.

Applied early-warning systems have generally ignored the problems associated with autocorrelation inherent in the econometric model. As a result, the estimations produce incorrect *t*-statistics.[4]

---

[2] The two year forecast horizon is used, for example, in Kaminsky, Lizondo, and Reinhart (1998) and Berg and Pattillo (1999). Of course, private sector/market-oriented models are much more interested in short-run forecasts. The issue is discussed in Berg, Borensztein, Milesi-Ferretti, and Pattillo (1999). For a recent survey of the crisis early warning system literature, see Abiad (2003).

[3] Poirier and Ruud (1988) and Gourieroux, Monfort, and Trognon (1984) establish this under fairly general conditions; later in this paper, we briefly discuss their application to our case. See Estrella and Rodriquez (1998) for a more extended discussion in a similar context.

[4] The previous published work of one of us is, of course, vulnerable to this criticism Berg and Pattillo (1999). Other papers that appear to contain estimates with the same error include Schnatz (1999); Goldstein, Kaminsky, and Reinhart (2000); Reinhart (2000); Bussiere and Fratzscher (2002); Komulainen and Lukkarila (2003); Sy (2003); and Leblang and Satyanath (2003). Using very different setups, Burkart and Coudert (2000) and Osband and Van Rijckeghem (2000) have related problems; the latter acknowledge and briefly discuss the issue. Some early-warning-system papers, particularly those used for analytic rather than predictive purposes, avoid the problem because they focus only

(continued…)

In this paper, we illustrate the problem and correct it in the context of the Berg and Pattillo (1999) (hereinafter referred to as BP) early-warning-system framework, a variant of which has been employed at the IMF (Berg, Borensztein, Milesi-Ferretti, and Pattillo, 1999). In our application, the dependent variable is highly serially-correlated, since if a crisis looms, say, 22 months ahead, it also looms 21 months ahead, and so on. Meanwhile, the levels of the explanatory variables, such as the degree of exchange rate overvaluation, tend to be similar from month to month. Thus, if a model fails to predict a crisis one month, it will likely fail to predict it the next.

To correct the standard error estimates, we follow Estrella and Rodriguez (1998) and compute a heteroskedasticity- and autocorrelation-corrected (HAC) covariance matrix along the lines of Hansen (1982). Finite-sample properties of such matrices depend on a variety of choices for which asymptotic theory provides little guidance.[5] We thus use Monte Carlo simulations to test the proposed procedure on simulated data that mimics key features of the BP data. For comparison, we also calculate bootstrap estimates of the standard errors.

We find that (1) the uncorrected probit results in substantial underestimates of the true standard errors, up to a factor of four for highly serially correlated explanatory variables; and (2) the HAC procedure produces accurate standard errors in the simulation. The bootstrap standard errors agree closely with the HAC estimates. Finally, we apply the HAC and bootstrap procedures. We find that (3) for the actual BP currency crisis data, the standard errors are greatly increased, though three out of the five variables remain significant at conventional significance levels; one becomes marginally insignificant, and the fifth loses significance.[6]

---

on the cross-section or because they look only at one-period-ahead forecasts (e.g., Fratzscher (2003), Eliasson and Kreuter (2001), Esquivel and Larrain (1998), Frankel and Rose (1996), Falcetti and Tudela (2002), Kamin, Schindler, and Samuel (2001)). The popular Kaminsky, Lizondo, and Reinhart (1998) indicators approach avoids this problem because it produces no standard errors at all.

[5] See den Haan and Levin (1997) for a review. The amount of information contained in a typical EWS sample may be smaller than the sample size would suggest. While the BP sample contains 5,025 observations, only 16 percent have a dependent variable with a value of 1.

[6] For the updated model described in Berg, Borensztein, Milesi-Ferretti, and Pattillo (1999), in which short-term debt/reserves replaces M2/reserves, all the variables remain at least marginally significant.

II. Empirical Framework

A. The Model

We briefly describe the Berg and Pattillo (1999) (BP) early-warning-system framework and report the estimated results. This model has displayed a reasonable degree of success in predicting crises, both in-sample and out-of-sample, and in an updated form has been used internally within the IMF, along with other models, since 1999. As discussed in the introduction, the problems we discuss are shared with a number of other EWS models.[7]

BP define $C_{i,t}=1$ if there is a currency crisis in period *t* for country i. A crisis occurs, by definition, when both the exchange rate and reserves decline by more than a critical level. We are interested in assessing the likelihood of a crisis sometime in the near future, rather than predicting the exact month of a crisis. Given the interest in predicting crises several months ahead, and the view that vulnerability but not exact time may be predictable, define $c24_{i,t} = 1$ if there is a crisis sometime within the next 24 months, 0 otherwise. Thus,

$$C24_{i,t} = 1 \quad iff \quad \sum_{r=1}^{24} C_{i,t+r} > 0 \qquad (1)$$

Further, the probability of a crisis in the next two years is a function of a set of variables, X:

$$P(C24_{i,t} = 1) = F(X_{i,t}\beta) \qquad (2)$$

where F is the cumulative normal distribution. There are five variables included in $X_{i,t}$: a measure of real exchange rate overvaluation, the 12-month[8] change in reserves, the 12-month change in exports, the four-quarter moving average of the current account balance as a share of GDP, and the ratio of M2 to reserves. BP use this model to estimate the 24-

[7] For a more extensive discussion of the role of this and other models in the Fund, see Berg, Borensztein, Milesi-Ferretti, and Pattillo (1999). For an out-of-sample evaluation of various crisis prediction models in practice, see Berg, Borensztein, and Pattillo (2004).

[8] In Kaminsky, Lizondo and Reinhart (1998), Berg and Pattillo (1999), and Berg, Borensztein, Milesi-Ferretti, and Pattillo (1999), the variables in $X_{i,t}$ are transformed into percentile terms, whereby each is measured in terms of percentiles of country i's distribution for the underlying variable.

month probability of a currency crisis over an unbalanced panel of 304 monthly observations (January 1970 to April 1995) from each of 23 countries (Table 1, columns 1 to 3).[9]

B. Serial Correlation Problem

The construction of the C24 variable induces serial correlation in the errors. Whenever C24 takes a value of 1, it does so for 24 consecutive periods, which implies the errors in the model are serially correlated for 24 periods. There is also substantial serial correlation in the right-hand-side variables. This induces a similar pattern of serial correlation in the residuals from the probit estimation, resulting in inconsistent estimates of the probit standard errors.

Figure 1 shows the value of C24, the predicted probabilities from the model, and the five right-hand-side variables for Argentina, the first country in the sample (alphabetically). It shows the high degree of persistence of both the left and most of the right-hand-side variables in equation 1.[10]

C. Heteroskedasicity- and Autocorrelation-Corrected (HAC) Estimates of Standard Errors

Following Estrella and Rodrigues (1998), we construct the usual probit (quasi-maximum likelihood (QML)) estimates of the beta coefficients as well as HAC estimates of the covariance matrix. Let $F(\cdot)$ and $f(\cdot)$ denote the cumulative normal and standard normal distributions respectively, with $C24_t$ and $\beta' X_t$ defined in equation 1 and 2 above.

---

[9] Results for the updated model used internally at the IMF and described in Berg, Borensztein, Milesi-Ferretti, and Pattillo (1999) (called the "DCSD" model for historical reasons) are reported in the second panel of Table 1. There are two differences between the models: the DCSD model uses short-term debt/reserves instead of M2/reserves and (2) the sample is smaller, dropping one country (Taiwan, Republic of China), and starting in 1985 instead of 1970 but continuing through 1997.

[10] Four of the variables are highly persistent. Coefficients in an AR1 representation of these series from a pooled regression are above 0.9 for four of the variables, with only export growth notably less persistent, with an AR1 coefficient of 0.57. Im-Pesaran-Shinn (2002) tests in heterogeneous panels nonetheless allow a rejection of the null of a unit root for all the variables except the current account/GDP, where the t_bar statistic has a value of -1.1 (the critical value at the 10 percent level is -1.75). We are comfortable assuming stationarity even here, on the grounds that it is *a priori* reasonable to suppose that the current account as a share of GDP does not contain a stochastic trend. Moreover, the time series is reasonably short for these purposes at about 25 years, implying low power to reject stationarity.

Consider the following log-likelihood function, which would be appropriate if the observations were iid:

$$\log L = \sum_{t=1}^{T} \left\{ c24_t F_t + (1 - c24_t)(1 - F_t) \right\} \tag{3}$$

Maximizing this log-likelihood function produces the usual probit results as QML estimates. The first-order conditions for the usual estimates of β are:

$$\sum_{t=1}^{T} h_t = 0$$

where

$$h_t = \frac{[C24_t - F(\beta' X_t)] f(\beta' X_t) X_t}{F(\beta' X_t)[1 - F(\beta' X_t)]} \tag{4}$$

The resulting estimates of β, denoted $\hat{\beta}^{QML}$, are the standard probit estimates. They are consistent in the presence of serial correlation.[11]

The maximum likelihood estimates of the variance-covariance matrix for the estimated coefficient estimates V_{β}^{QML} is equal to

$-H_0^{-1}$, *where* $H_0 = \sum_{t=1}^{T} \frac{-f_t^2 X_t X_t'}{F_t(1 - F_t)}$. As mentioned above, in general this estimate will be inconsistent in the presence of serial correlation of the C24 variable and the X variables.

An alternative estimator of the V_β is the HAC estimator, associated with Generalized Method of Moments (GMM) estimation techniques. This estimator is built from the sample autocovariances of h (from equation 4):

[11] Sufficient conditions are (1) that the errors in the probit relationship are homoskedastic and well-behaved and (2) that the distribution of $X_{i,t}$ is stationary and ergodic, as discussed in Wooldridge (1994). We assume this here, with stationarity discussed in footnote 10.

$$\hat{\Omega}_j = \frac{1}{T}\sum_{t=j+1}^{T} h_t h_{t-j}^{'} \qquad (5)$$

Following Hansen (1982), we construct a family of estimators of the covariance of h:

$$\hat{S} = \hat{\Omega}_0 + \sum_{j=1}^{m} \lambda_j (\hat{\Omega}_j + \hat{\Omega}_j^{'}) \qquad (6)$$

where $\lambda_j = 1$ in Hansen (1982) and $\lambda_j = 1 - j/(m+1)$ in Newey and West (1987). The parameter m represents the maximum number of lags across which the h_t terms may be correlated. If m grows with the sample size, this matrix is a consistent estimator of the covariance matrix of the orthogonality conditions.

As shown in Estrella and Rodrigues (1998) and following Hansen (1982), the HAC heteroskedasticity- and autocorrelation-corrected estimate of the variance-covariance matrix of the QML estimator of the coefficients is

$$V_{\beta}^{HAC} = H_0^{-1} \hat{S} H_0^{-1} \qquad (7)$$

We follow Hansen (1982) in assuming λ_j =1; that is we assume an equal weighting of the interactions between current observations and past observations. We also set m=30, which means the covariance estimator only incorporates the interactions between the current and past thirty observations for each *t* (remember from equations 1 and 2 that the overlapping forecasts are 24 periods ahead).[12]

D. Bootstrap Estimates of Standard Errors

An alternative estimate of V_β is readily available through bootstrap methods. The basic idea is simple. (i) We draw, with replacement, from the set of actual data used to estimate equation 1. That is, we start with an empty data set. We then pick a country at random and include the actual data from that country to the data set. We repeat, with replacement, as many times as there are countries in the sample. (2) Estimate $\hat{\beta}^{QML}$ as described

[12] We discuss below the sensitivity of our results to variants such as using Newey-West weighting, different lag lengths, and different sample sizes. The structure of C24 suggests that the serial correlation for that variable should be constant for 24 periods then drop off to nothing, providing some a priori justification for our choice of kernel and for setting m equal to a number slightly above 24.

above. Call this estimate $\hat{\beta}^{BS}(j)$. (3) Repeat j times (we typically generated 500 estimates). Then estimate the bootstrap variance of $\hat{\beta}$, V_{β}^{BS}, as the variance of $\hat{\beta}^{BS}(j)$ across the j samples.[13] Under certain conditions, this is a consistent estimator with acceptable small-sample properties.

III. Monte Carlo Simulations

Before applying our HAC and bootstrap variance estimators to the actual BP data, we use Monte Carlo simulations to demonstrate how inaccurate the standard probit standard errors are in this situation. We then ask whether the HAC and bootstrap estimators are better. In order to carry out these tests, we generate data similar to that used by BP and apply the various estimators. By repeating this process many times, we can observe the actual variance (across simulation runs) of the estimators, and compare this to the estimates of the variance produced by the regressions.

We first generate a panel data set of independent variables with autocorrelation properties similar to those in the BP data. Specifically, for each country i we generated five independent variables $X_{i,t}$ as AR1 processes with normal errors, with the same first-order autocorrelation coefficients as in the five variables in the BP data. We then construct a crisis variable $C_{i,t}$ and the resulting C24 variable as follows:

$$C_{i,t} = 1 \quad iff \quad F(X_{i,t}\gamma) > -\varepsilon_{i,t}\, for\, \varepsilon_{i,t} \sim N(0,1) \qquad (8)$$

$$C24_{i,t} = 1 \quad iff \quad \sum_{j=0}^{23} C_{i,t+j} > 0 \qquad (9)$$

We then estimate the BP model described in equation 2 using both the standard probit methodology and also calculating the HAC and bootstrap variance-covariance matrices described above. We repeat 500 times. We can then calculate the standard probit results, that is the mean values across the 500 runs of the estimated betas, $\bar{\hat{\beta}}^{QML}$, the mean values of the estimated standard errors, $\left(\bar{V}_{\beta}^{QML}\right)^{1/2}$ and the mean associated t-statistics. We

[13] See Efron and Tibshirani (1994) for an introduction. Note that the bootstrap technique could be used to calculate confidence intervals directly, which would avoid requiring the assumption that $\hat{\beta}$ is distributed normally. Instead, we calculate the bootstrap standard errors and implied t-statistics because we are using the bootstrap estimates primarily to check our HAC results.

report these in columns 1, 3 and 4 of Table 2. [14] For comparison, column 2 shows the standard deviation across simulations of $\hat{\beta}$. A good standard error estimate for β would be close to the observed standard deviation of β. In fact, the actual variance of β across runs is much higher than the average estimated standard error, by 50 percent for the relatively non-persistent export growth variable but by a factor of above 4 for the most persistent variables.

The accuracy of the estimated standard errors is soundly rejected statistically. We take as the null hypothesis that $\beta = \bar{\hat{\beta}}^{QML}$ and that the standard error estimates are correct. If true, we should find this hypothesis rejected at the 5 percent significance level in about 5 percent of the simulations. The fifth column of Table 2 shows the fraction of the simulations in which this hypothesis is in fact rejected, and in parentheses the probability of finding rejections this often if the standard error estimates are correct. All these p-values are below 0.01.

How do the alternative standard error estimates fare? Table 2 reports the mean values of the HAC standard errors associated with $\bar{V}_{\beta}^{HAC}$ and the associated t statistics for the hypothesis $\beta = 0$. These estimated standard errors match the observed standard deviation of $\hat{\beta}$ quite closely (within ± 8 percent). For all but one of the variables, we cannot reject the hypothesis that the estimated standard errors are correct at conventional measures. Even for the fifth variable, export growth, the standard error is qualitatively close though statistically not perfect. The bootstrap estimates are similar to the HAC estimates, though slightly more accurate.[15]

The HAC corrections do not create major problems in situations where they are not necessary. For similar simulations but for serially uncorrelated X variables, the maximum

---

[14] Note that we have not generated our simulated C24 variable based on its relationship to $F(X\beta)$, but rather indirectly through an assumed relationship between C and $F(X\gamma)$. That is, data generating process requires an assumption on γ, not β. We do not know the theoretical relationship between γ and β, so we do not know what the value of β should result from the monte carlo estimation. Thus, we cannot confirm directly that our probit estimates of β, $\hat{\beta}^{QML}$, are consistent. We rely on the theorems discussed in footnote 11. In practice, in setting up the monte carlo we initially experiment with different values of γ until we find those that result in estimated values of β that are close to the values observed in BP.

[15] The tests reported in column 5 are for two-tailed confidence intervals. For one-tailed tests (i.e. where the alternative hypothesis is that β>0), the HAC and bootstrap standard errors are somewhat better still and the hypothesis that the tests have the correct size cannot be rejected at standard significance levels.

likelihood estimates of the standard errors are fine, but so are the HAC and bootstrap standard errors (Table 3).[16]

The HAC estimators require large sample sizes in our context. Figure 2 shows the mean values across simulations of the estimated standard errors , $\left(\overline{V}_{\beta}^{QML}\right)^{1/2}$, divided by the standard deviation $\hat{\beta}$, for simulations identical to that presented in Table 2 but for different sample sizes. With samples of about 3000, such as those used to estimate the DCSD EWS model, the HAC standard errors work about as well as with the BP sample size presented in Table 2. As sample sizes fall to about 1500 or so, though, the standard errors are off by amounts that would likely be unacceptable economically, with standard errors up to 30 percent too small.[17]

IV. Estimation of Currency Crisis Data

We now apply are alternative estimators to the actual data used in BP. Table 1 columns 4 through 7 show the estimation results for the original uncorrected probit model as well as the HAC and bootstrap-based standard errors. As expected from the simulations, the standard errors are much higher, particularly for the most persistent variables. Three of the five variables remain significant at the 5 percent level however. One more (export growth) is now significant only at the 10 percent level, while M2/Reserves is not significant. Note that, as above, the bootstrap and the HAC estimates agree fairly closely with each other.[18]

V. Conclusion

In this paper, we have documented and addressed the problems caused by serial correlation in the context of crisis early-warning-system models estimated as panel probit regressions—the estimates of standard errors tend to be much too low. The problem is quite general and occurs whenever the model forecasts several periods ahead. The need to

---

[16] For the simulations in Table 3, none of the independent variables are serially correlated. Note that the estimated maximum-likelihood standard error for the constant is still incorrect.

[17] We tried using Newey-West standard errors, but found that this method worked more poorly in our framework. We also experimented with different lag lengths (m in equation (6)) and found that increasing above 30 did not help, while going much smaller (say 15) resulted in poorer estimates. Results are available from the authors on request.

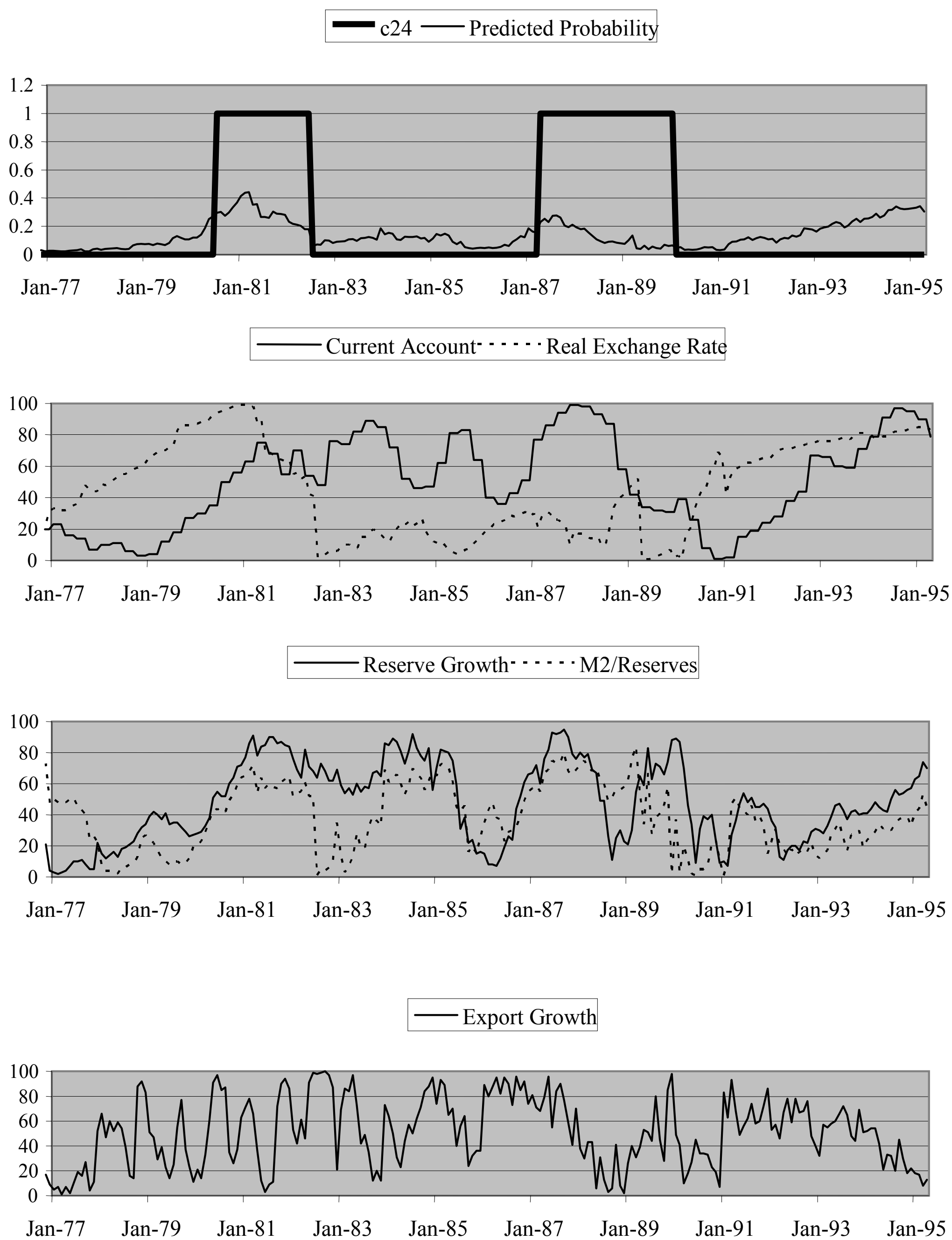
[18] For the DCSD model presented in the bottom panel of Table 1, all the variables remain significant, though export growth marginally so.

provide up-to-date forecasts over substantial horizons implies that most applied models will have this feature.

We have demonstrated that applying a standard maximum-likelihood probit technique to data such as that used in BP results in substantial underestimation of the size of the standard errors, by as much as a factor of four for highly persistent independent variables. We have also demonstrated through Monte Carlo simulation that HAC standard errors are good estimates of the actual standard errors. We confirmed this by finding that a quite different bootstrap technique gives similar results. Applying these techniques to the BP regressions, we find that most of the variables remain statistically significant, though substantially less so than the original probit estimates suggested.

This paper could be a footnote: because of overlapping forecasts and the resulting uncorrectable serial correlation, GMM is the appropriate estimation technique for the panel probits used in BP and similar models. We have expanded it for several reasons. First, a number of papers, particularly by practitioners, have been making the same mistake as BP in running simple probit regressions. In part, this may be because few packages can automatically produce a HAC estimator for a nonlinear model, such as a probit in a panel setting. Second, theory does not tell us much about how serious the problem is in practice or whether the asymptotically correct solutions will work well in a particular sample. Third, the new results with respect to the BP and related models may themselves be of interest.

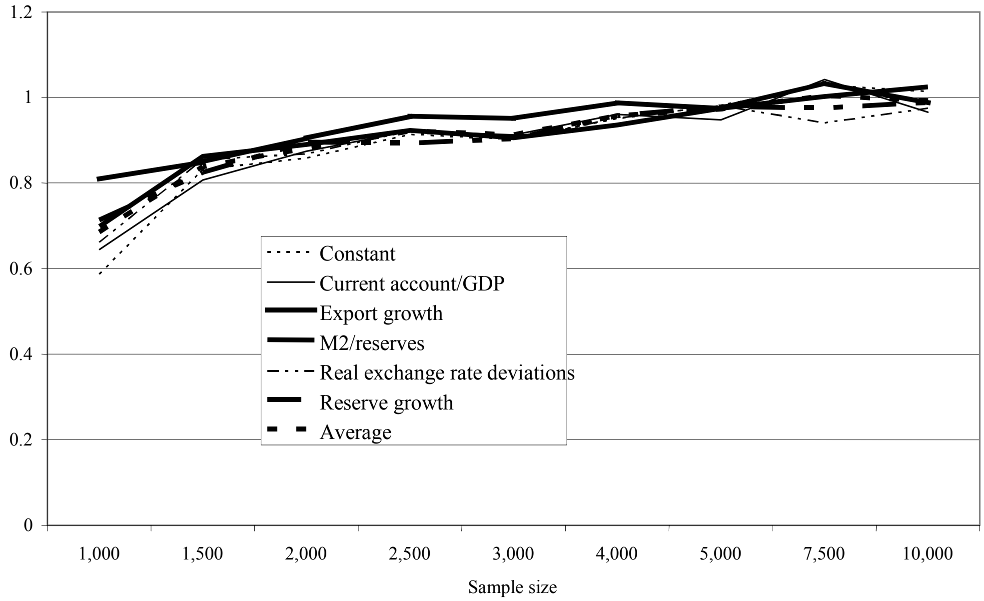
Figure 1. Argentina: C24, Predicted Probabilities, and Independent Variables



Sources: IMF, *International Financial Statistics*; and IMF staff estimates.

Notes:
C24 takes a value of 1 when there is a crisis sometime in the next 24 months, 0 otherwise.
Predicted probability is the probability of a currency crisis sometime in the next 24 months, following Berg and Pattillo (1999).
Other variables are transformed into percentiles. Each is measured in terms of percentiles of Argentina's distribution for the underlying variable.

Figure 2. Mean of Estimated Standard Errors / Standard Deviation of Estimated Coefficient



Source: Authors' calculations

Table 1. Probit Early-Warning-System (EWS) Models

	Original Estimates			Alternate Standard Errors			
		Maximum Likelihood		HAC		Bootstrap	
	Coefficient	std. err.	t-statistic	std. err.	t-statistic	std. err.	t-statistic
Variable [1/]	(1)	(2)	(3)	(4)	(5)	(6)	(7)
BP Model (from Berg and Pattillo (1999))							
Sample: 5025 observations (23 countries, 1970:1 to 1995:4) of which 813 with dependent variable = 1							
Constant	-2.6389	0.089854	-29.37	0.316534	-8.34	0.397478	-6.64
Current account/GDP	0.0083	0.000872	9.50	0.002775	2.99	0.003515	2.36
Export growth	0.0030	0.000817	3.65	0.001675	1.78	0.002063	1.45
Real exchange rate deviations	0.0108	0.000797	13.50	0.003238	3.32	0.003881	2.77
Reserve growth	0.0060	0.000961	6.20	0.002515	2.37	0.003189	1.87
M2/reserves	0.0025	0.000880	2.80	0.003665	0.67	0.004095	0.60
DCSD Model (Based on Berg, Borensztein, Milesi-Ferretti, and Pattillo (1999))							
Sample: 3034 observations (22 countries, 1985:12 to 1997:12) of which 518 with dependent variable = 1							
Constant	-3.8756	0.151446	-25.59	0.511906	-7.57		
Current account/GDP	0.0158	0.001235	12.83	0.004122	3.84		
Export growth	0.0045	0.001253	3.59	0.002847	1.58		
Real exchange rate deviations	0.0121	0.001188	10.19	0.004375	2.77		
Reserve growth	0.0089	0.001189	7.47	0.003497	2.54		
Short-term external debt/reserves	0.0149	0.001155	12.91	0.004480	3.33		

Source: Authors' calculations.

Note: The abbreviation "std. err." denotes standard errors.

1/ All coefficients should be positive since variables such as reserve growth, export growth, and real exchange rate deviations from trend have been multiplied by -1.

Table 2. Comparison of Standard Error Estimates: Monte Carlo Simulation Results for BP-Like Data [1/]
(500 Monte Carlo runs, 200 bootstrap draws each)

			Standard Errors								
			Maximum Likelihood			HAC			Bootstrap		
Variable	$\hat{\beta}$ mean (1)	$\hat{\beta}_{s.d.}$ (2)	Mean of $\hat{\sigma}$ (3)	$\bar{t}$ (4)	%reject (p value) (5)	Mean of $\hat{\sigma}$ (6)	$\bar{t}$ (7)	%reject (p value) (8)	Mean of $\hat{\sigma}$ (9)	$\bar{t}$ (10)	%reject (p value) (11)
Constant	-2.51	0.33	0.10	-25.4	0.53 (0.00)	0.33	- 7.7	0.06 (0.31)	0.33	-7.7	0.06 (0.54)
Current Account/GDP	0.81	0.32	0.07	11.1	0.63 (0.00)	0.29	2.8	0.08 (0.01)	0.01	591.8	0.07 (0.07)
Export Growth	0.34	0.13	0.09	3.8	0.18 (0.00)	0.14	2.4	0.05 (0.84)	0.27	464.6	0.06 (0.31)
Real exchange rate	0.90	0.27	0.06	14.1	0.65 (0.00)	0.25	3.6	0.07 (0.02)	0.06	560.0	0.07 (0.02)
Reserve growth	0.64	0.28	0.09	7.3	0.57 (0.00)	0.29	2.2	0.04 (0.22)	0.67	485.1	0.05 (0.84)
M2/reserves	0.26	0.31	0.08	3.3	0.61 (0.00)	0.30	0.9	0.06 (0.54)	0.21	539.0	0.05 (0.68)

Source: Authors' calculations

Notes:
There are 5,014 observations (of which 16 percent with dependent variable = 1) in 23 countries.
Columns (1), (3), and (4) present the mean across the 500 simulation runs of the estimated values of β, σ at t respectively, estimated as standard probit regressions. Column (2) presents the standard deviation of the estimated β across the 500 runs. Column 5 shows the fraction of the runs in which the hypothesis $\beta = \bar{\hat{\beta}}$ (i.e. that β = its average estimated value across the 500 runs) can be rejected in favor of the hypothesis that $\beta \neq \bar{\hat{\beta}}$. In parentheses of column (5) is the p-value that this hypothesis would be rejected this often if the standard error estimates were correct and $\hat{\beta}$ were distributed normally. Columns (6) through (8) present the same results for the HAC standard error estimates and (9) through (11) for the bootstrap standard error estimates.

1/ BP denotes Berg and Pattillo (1999)

Table 3. Comparison of Standard Error Estimates: Monte Carlo Simulation Results for BP-Like Data,[1/] No Serial Correlation in Independent Variables
(500 Monte Carlo runs, 200 bootstrap draws each)

			Standard Errors								
			Maximum Likelihood			HAC			Bootstrap		
	$\hat{\beta}$ mean	$\hat{\beta}_{s.d.}$	Mean of $\hat{\sigma}$	$\bar{t}$	%reject (p value)	Mean of $\hat{\sigma}$	$\bar{t}$	%reject (p value)	Mean of $\hat{\sigma}$	$\bar{t}$	%reject (p value)
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Constant	-2.35	0.21	0.12	-19.6	0.24 (0.00)	0.21	-11.5	0.06 (0.31)	0.20	-12.1	0.07 (0.10)
Current Account/GDP	0.69	0.20	0.20	3.4	0.04 (0.41)	0.19	3.7	0.05 (1.00)	0.20	3.7	0.05 (0.84)
Export Growth	0.39	0.10	0.09	4.1	0.06 (0.22)	0.09	4.2	0.08 (0.01)	0.09	4.3	0.08 (0.00)
Real exchange rate	0.53	0.19	0.18	3.0	0.05 (0.84)	0.17	3.1	0.07 (0.04)	0.17	3.1	0.10 (0.00)
Reserve growth	0.51	0.17	0.17	3.0	0.05 (0.84)	0.16	3.2	0.06 (0.22)	0.16	3.2	0.08 (0.00)
M2/reserves	0.51	0.19	0.19	2.7	0.05 (0.84)	0.18	2.9	0.08 (0.00)	0.18	2.9	0.09 (0.00)

Source: Authors' calculations

Notes:
There are 5,014 observations (of which 16 percent with dependent variable = 1) in 23 countries.
None of the five independent variables excluding the constant contain serial correlation.
Columns (1), (3), and (4) present the mean across the 500 simulation runs of the estimated values of β, σ at t respectively, estimated as standard probit regressions. Column (2) presents the standard deviation of the estimated β across the 500 runs. Column 5 shows the fraction of the runs in which the hypothesis $\beta = \bar{\hat{\beta}}$ (i.e. that β = its average estimated value across the 500 runs) can be rejected in favor of the hypothesis that $\beta \neq \bar{\hat{\beta}}$. In parentheses of column (5) is the p-value that this hypothesis would be rejected this often if the standard error estimates were correct and $\hat{\beta}$ were distributed normally. Columns (6) through (8) present the same results for the HAC standard error estimates and (9) through (11) for the bootstrap standard error estimates.
1/ BP denotes Berg and Pattillo (1999)

REFERENCES

Abiad, Abdul, 2003, "Early Warning Systems: A Survey and a Regime-Switching Approach," Working Paper No. 03/32 (Washington: International Monetary Fund).

Berg, Andrew, Eduardo Borensztein, Gian Maria Milesi-Ferretti, and Catherine Pattillo, 1999, "Anticipating Balance of Payments Crises: The Role of Early Warning Systems," IMF Occasional Paper 186 (Washington, D.C.: International Monetary Fund).

———, and Catherine Pattillo, 1999, "Predicting Currency Crises: The Indicators Approach and an Alternative," *Journal of International Money and Finance,* Vol. 18, Issue 4 (August), pp. 561–86.

———, Eduardo Borensztein, and Catherine Pattillo, 2004, "Assessing Early Warning Systems: How Have They Worked in Practice?," Working Paper, forthcoming (Washington, D.C.: International Monetary Fund).

Burkart, Olivier, and Virginie Coudert, 2000, "Leading Indicators of Currency Crises in Emerging Economies," Notes d'Etudes et de Recherche # 74 (Paris: Banque de France).

Bussiere, Matthieu, and Marcel Fratzscher, 2002, "Towards a New Early Warning System," Working Paper No. 145 (Frankfurt: European Central Bank).

den Haan, Wouter J., and Andrew T. Levin, 1997, "A Practitioner's Guide to Robust Covariance Matrix Estimation." in *Robust Inference*, G. S. Maddala and C. R. Rao eds., pp. 299–342 (Handbook of Statistics series, Vol. 15. Amsterdam; New York and Oxford: Elsevier Science North-Holland).

Efron, Bradley, and Robert J. Tibshirani, 1994, *An Introduction to the Bootstrap* (Boca Raton: Chapman and Hall).

Eliasson, Ann-Charlotte, and Christof Krueter, 2001, "On Crisis Models: An Alternative Crisis Definition," Working Paper (Frankfurt: Deutsche Bank Research).

Estrella, Arturo, and Anthony Rodrigues, 1998, "Consistent Covariance Matrix Estimation in Probit Models with Autocorrelated Errors," Staff Report No. 39 (New York: Federal Reserve Bank of New York).

Frankel, Jeffrey A., and Andrew K. Rose, 1996, "Currency Crashes in Emerging Markets: An Empirical Treatment," *Journal of International Economics,* Vol. 41, Issue 3–4 (November), pp. 351–66.

Fratzscher, Marcel, 2003, "On Currency Crises and Contagion," *International Journal of Finance and Economics,* Vol. 8, Issue 2 (April), pp. 109–29.

Goldstein, Morris, Graciela L. Kaminsky, and Carmen M. Reinhart, 2000, *Assessing Financial Vulnerability: An Early Warning System for Emerging Markets* (Washington, D.C.: Institute for International Economics).

Gourieroux, Christian, Alain Monfort, and Alain Trognon, 1984, "Estimation and Test in Probit Models with Serial Correlation." in *Alternative Approaches to Time Series Analysis*, J.P. Florens, M. Mouchart, J.P. Raoult and L. Simar eds., pp. 169–209 (Brussels: Facultes Universitaires Saint-Louis).

Hansen, Lars Peter, 1982, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica,* Vol. 50, Issue 4 (July), pp. 1029–54.

Im, Kyung So, M. Hashem Pesaran, and Yongcheol Shin, 2003, "Testing for Unit Roots in Heterogeneous Panels," *Journal of Econometrics,* Vol. 115, Issue 1 (July), pp. 53–74.

Kamin, Steven B., John W. Schindler, and Shawna L. Samuel, 2001, "The Contribution of Domestic and External Factors to Emerging Market Devaluation Crises: An Early Warning Systems Approach," International Finance Discussion Paper No. 711 (Washington, D.C.: Board of Governors of the Federal Reserve System).

Kaminsky, Graciela, Saúl Lizondo, and Carmen M. Reinhart, 1998, "Leading Indicators of Currency Crises," *International Monetary Fund Staff Papers,* Vol. 45, Issue 1 (March), pp. 1–48.

Komulainen, Tuomas, and Johanna Lukkarila, 2003, "What Drives Financial Crises in Emerging Markets," *Emerging Markets Review,* Vol. 4, pp. 248–72.

Leblang, David, and Shankar Satyanath, 2003, "Politics as an Early Warning for Currency Crises," Paper to be presented at the 2003 Meetings of the American Political Science Association (Boulder: University of Colorado).

Newey, Whitney K., and Kenneth D. West, 1987, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica,* Vol. 55, Issue 3 (May), pp. 703–08.

Osband, Kent, and Caroline Van Rijckeghem, 2000, "Safety from Currency Crashes," *IMF Staff Papers,* Vol. 47, Issue 2 pp. 238–58.

Poirier, Dale J., and Paul A. Ruud, 1988, "Probit with Dependent Observations," *Review of Economic Studies,* Vol. 55, Issue 4 (October), pp. 593–614.

Reinhart, Carmen M., 2000, "Default, Currency Crises, and Sovereign Credit Ratings," NBER Working Paper No. 8738 (Cambridge, MA: National Bureau of Economic Research).

Schnatz, Bernd, 1999, "The Sudden Freeze of the Asian Miracle: The Role of Macroeconomic Fundamentals," *Asia Pacific Journal of Finance,* Vol. 2, pp. 1–19.

Sy, Amadou, 2003, "Rating the Ratings Agencies: Anticipating Currency Crises or Debt Crises," Working Paper No. 03/122 (Washington: International Monetary Fund).

Wooldridge, Jeffrey M., 1994, "Estimation and Inference for Dependent Processes." in *Handbook of Econometrics*, Robert F. Engle and Daniel L. McFadden eds. (Amsterdam; London and New York: Elsevier North-Holland).