Available online at www.sciencedirect.com
ScienceDirect
Computational Statistics & Data Analysis 51 (2006) 1420–1441
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

Early warning systems for sovereign debt crises: The role of heterogeneity

Ana-Maria Fuertes*, Elena Kalotychou*

Faculty of Finance, Cass Business School, City University, 106 Bunhill Row, London EC1Y 8TZ, UK

Received 23 March 2006; received in revised form 17 August 2006; accepted 18 August 2006
Available online 12 September 2006

Abstract

Sovereign default models that differ in their treatment of unobservable country, regional and time heterogeneities are systematically compared. The analysis is based on annual data over the 1983–2002 period for 96 developing economies. Inference-based criteria and parameter plausibility overwhelmingly favour more complex models that allow the link between the probability response and the fundamentals to vary over time and across countries. However, out-of-sample forecast evaluation using several loss functions and equal-predictive-ability tests suggests that simplicity beats complexity. Parsimonious pooled logit models produce the most accurate sovereign default forecasts and outperform the naive benchmarks.

Keywords: Credit risk; Default probability; Emerging markets; Loss function; Panel logit; Predictive performance; Unobserved heterogeneity

1. Introduction

The sovereign debt crises of recent decades have emphasized the importance of credit risk prediction. Financial institutions use default probabilities to price loans and bonds, to determine adequate concentration limits and as inputs for value-at-risk analyses. Furthermore, the new Basel Accord allows banks to use internal ratings and default probabilities to set their regulatory capital. The available studies can be grouped into three broad types. One group exploits option pricing models to obtain the implied default probability while another seeks to explain the default probability using structural models or panel discrete-choice models. This study belongs to the latter group. A third research area focuses on explicitly generating default probabilities from credit ratings. This is the approach adopted in Fuertes and Kalotychou (2006a) where the finite-sample properties of different credit migration estimators are investigated by bootstrap simulations.

A large empirical literature analyses the determinants of sovereign default and the results are quite mixed. Some studies find that liquidity or global business cycle indicators are crucial. The evidence on the importance of structural economic conditions also varies across studies. Most researchers estimate logit models using pooled data for a large number of countries although the validity of the underlying homogeneity assumption has been questioned. McFadden et al. (1985) and Hajivassiliou (1987) argue that the link between debt repayment performance and macroeconomic attributes is likely to vary across countries and over time. For countries with fewer capital controls or more open to

* Corresponding authors. Tel.: +44 20 7040 8647; fax: +44 20 7040 8881.
E-mail addresses: a.fuertes@city.ac.uk (A.-M. Fuertes), e.kalotychou@city.ac.uk (E. Kalotychou).

0167-9473/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2006.08.023

global trade, external economic activity signals may have more predictive power. Moreover, qualitative idiosyncrasies (colonial histories, types of government and religious institutions) may explain why country A defaults but country B services its debt while exhibiting similar economic fundamentals and debt structures. As Shleifer (2003; p. 5) succinctly puts it: "Sovereign debt markets could not be more different".

The few studies that control for *country* heterogeneity adopt either fixed or random effects models (Detragiache and Spilimbergo, 2001; Li, 1992; Oral et al., 1992). Evidence from the currency crisis literature suggests that the relevant heterogeneity is *regional* rather than country specific (Burkart and Coudert, 2002; Staikouras, 2004; Kalotychou and Staikouras, 2005). For instance, different key leading indicators of currency crises have been identified for Asia and Latin America. In sovereign default studies, regional differences have been captured using dummy variables (Feder et al., 1981). *Time* heterogeneity may reflect changing world conditions such as the business cycle and the development of international capital markets or the changing nature of emerging markets themselves. To control for such time effects that are assumed to be common across countries, some studies include year-dummy variables (Aylward and Thorne, 1998) whereas others use global macroeconomic indicators such as OECD growth (Lee, 1991; Detragiache and Spilimbergo, 2001).

The issue of whether controlling for country, regional or time heterogeneity helps to improve the forecast power of sovereign default models is relevant to regulators, practitioners and the rating agencies, all of which are mostly interested in the *when* rather than the *why* question of default. Regulators rely on default forecasts to monitor the financial health of banks, pension funds and other financial institutions that include sovereign debt in their portfolios. Practitioners feed such forecasts into theoretical or simulation models for pricing sovereign debt. Moreover, default probability predictions are used to test various hypotheses such as that country risk is priced in stock returns and borrowing costs. Finally, Early Warning Systems (EWSs) are recognized as a potentially fruitful complement to the broader analysis and judgement of decision-makers for identifying looming debt crises (Berg et al., 1999; Fuertes and Kalotychou, 2006b).

Against this background, it is surprising that forecasting issues have received scant attention in the literature. Most studies compare sovereign default models on the basis of their in-sample forecasts (Hajivassiliou, 1987; Detragiache and Spilimbergo, 2001). A few studies conduct out-of-sample evaluation but the forecasts are based on parameters estimated once and are limited to a 1- or 2-year holdout period. The forecast accuracy metrics typically used are Type I, Type II and overall error rates (Feder et al., 1981; Manasse et al., 2003; Oka, 2003; Peter, 2002). Furthermore, the few studies that provide out-of-sample predictions do not confront them with simple benchmarks such as those from random walk type models. This is particularly important in the present context due to the persistence in debt-servicing behaviour.

This paper contributes to the literature in two respects. First, to investigate the importance of unobserved heterogeneities in modelling sovereign default, it considers a wide range of logit models that differ in how they treat country, regional and time effects. The models' ability to describe the data generating process is gauged on the basis of statistical tests, information criteria and parameter plausibility. By characterizing the heterogeneities in different ways we seek to assess whether these unobserved effects are genuine or merely an artefact of misspecification. Some of the parameterizations, such as the random coefficients model with time-dependent and country-specific slopes and the models that allow for region-specific or time-specific slopes, have not yet been utilized in the present context. In addition, three novel global indicators are included to control for time effects in the likelihood of sovereign default. Second, a comprehensive forecasting analysis is conducted. A 12-year window is rolled forward to generate out-of-sample forecasts sequentially over 5 years. A battery of tests is run to assess the importance of heterogeneity in forecasting. For this purpose, several forecast accuracy measures are adopted: probability scores as well as metrics that allow for asymmetric misclassification costs. The tests are applied not only over the entire 5-year holdout sample but also over a positive-directional-change subset in order to gauge the models' ability to predict new (as opposed to ongoing) defaults. Various uninformative benchmarks are considered including random walk type models and the naive models implicit in the Pesaran and Timmermann (1992) and Donkers and Melenberg (2002) tests.

The statistical tests and model selection criteria indicate that the more complex specifications describe the data better. Unobserved heterogeneity across countries, regions and time is significant in explaining sovereign default. By contrast, the forecast horse race suggests that the parsimonious pooled logit model that imposes full homogeneity appears capable of yielding relatively good out-of-sample predictions and beating the benchmark models. Hence, our findings corroborate in this novel context the well-known limited relationship between in-sample fit and out-of-sample forecast performance. Simple variants of the pooled logit model that allow for fixed regional effects or time effects also