Credit Ratings in the Presence of Bailout: the Case of Mexican Subnational Government Debt

By

Fausto Hernández-Trillo* and Ricardo Smith-Ramírez
División de Economía
CIDE

ABSTRACT

Looking for an explanation for investment grades assigned to virtually bankrupt sub-national governments in LDCs, we study the determinants of bond ratings for municipalities in Mexico. Our data set includes ratings from three agencies: S&P, Fitch, and Moody's. To control for selectivity in the process of choosing an agency, we model the problem as a tri-variate self-selection process with ordinal responses. Additionally, in order to circumvent the estimation of multidimensional integrals, we implement a Monte Carlo Expectation Maximization (MCEM) algorithm. We find that not only financial but also political factors, such as number of voters and political party in power, are important and show evidence that the probability of bailout has a heavy weight in the rating process. Our outcomes question the purpose of rating sub-national debt in LDCs with a bailout tradition, since in those cases the market may assess the risk of sub-national entities as that of sovereign instruments.

JEL Classification: G18, G19, C3, H74

Key words: Credit Ratings, Bailout, Subnational Governments, Debt.

---

* Corresponding author. Address: Carretera México Toluca 3655; 01210 México D.F. MEXICO. Email: fausto.hernandez@cide.edu Phone: (52)(55) 57279839

1. INTRODUCTION

Bond ratings have existed for nearly a century. Debt issued by firms, sovereign countries and sub national governments (SNG[1]) are regularly rated in industrial countries (Cantor and Packer 1995, 1996). The rating history for less developed countries (LDCs) is shorter. International raters turned their attention to LDCs in the 1980s when agencies started rating LDC sovereign bonds as a reaction to several international debt crises. As a result, literature on grading SNGs and sovereign bonds in industrial countries abound, while for LDCs it is scarce.

Rating agencies have recently come under questioning in regard to the grading of LDCs. For example, in the January 5, 2004 issue of the *Wall Street Journal,* it was reported that credit ratings in China could merely be guesswork. In the case of sovereign credit ratings, there is a growing body of literature that casts doubt on its role, especially after the Asian and Argentinean crises of 1997-98 and 2001, respectively (e.g., Reinhart 2001 and 2002). Others have attempted to refine the measurement of risk (e.g. Remolona et. al. 2007 and Alfonso 2003). More recently, graders have been under question due to the delayed reaction in the US mortgage crisis[2]. In this paper, we target the rating technology used by agencies to grade SNGs in LDCs. By using data from the SNG bond market of Mexico, a country with a tradition of bailout, we analyze how political and financial factors are weighted in the construction of ratings. This case can be similar to most Latin American countries.

[1] A pioneering work for SNGs is Carleton and Lerner (1969).
[2] The Economist reported that the state of Ohio is suing a Credit Rater.

Latin American governments have a long tradition of bailing out SNGs; Bevilaqua (2000) documents this phenomenon for Brazil and Sanguinetti et al.(2000) does it for Argentina. A high bailout probability raises at least two issues: i) the adequacy of the bond rating process, and ii) its purpose. When establishing rating principles, many differences between industrial and LDC countries should be taken into account (Laulajainen, 1999). Typically, developing countries have serious institutional and legal shortcomings (see IADB, 1997). Most relevant, they are very centralized and most of them have just started fiscal decentralization reform, which in many cases has responded more to political pressure than to efficiency-enhancing purposes (see Díaz, 2006; Giugalle et al. 2001); they are more prone to financial crises, and market volatility is greater (see Bekaert and Campbell 1997); and law enforcement is deficient (La Porta and López-de-Silanes 2002). These characteristics are important when rating bonds in their local currencies (Mexican SNGs are not allowed to issue debt denominated in foreign currency); they call for different rating technologies than those used when rating industrial countries, where many of these shortcomings are not present.

Surprisingly, one of the largest states in Mexico (the State of Mexico) has been *continuously* bailed out since 1995 (virtually a bankrupt SNG) and has still been assigned an investment grade[3]. Likewise, Sanguinetti (2002) reports that one Argentinean provincial government (La Rioja) was bailed out several times previous to the 2001 crisis, and it still continues receiving an investment grading[4]. Since bond ratings are meant to indicate the likelihood of default (Bhatia 2002)[5], if SNGs are to be bailed out anytime they face

[3] Reported by *Bloomberg* on August 13, 2003 by Thomas Black. The grade assigned by Fitch is BBB.
[4] This author, among others, argues that the Argentinean crisis was in part due to fiscal permissiveness of SNGs governments in that country. For this reason, raters were questioned in Argentina.
[5] It has been shown that these agencies specialize in gathering and processing financial information and are certified by screening agents, who in turn are able to diversify their risky payoffs. In this setting raters solve,

financial problems, then their risk is passed on to the federal government. Therefore, SNG rates have eventually to become similar to those of sovereign debt. Is this happening in LDCs? If so, then the purpose of rating SNG debt may be arguable.

Rating agency behavior results are puzzling in LDCs, since (as pointed out before) they often give high rates to financially bankrupt SNGs. Then, what are agencies actually rating? Are they rating financial soundness or just probability of bailout? In this article, we try to answer these questions. Bailout events are most frequently the result of political negotiations. Therefore, we focus our analysis on the relevance (if any) that political factors have in the rating technology of three agencies: Standard and Poor's (S&P), Fitch, and Moody's. More specifically, we want to know whether number of voters and political party in power matter; given Mexico's long bailout tradition, we expect that they do. Additionally, we analyze how important financial issues are related to political factors. Finally, unlike previous studies, we study the determinants of choosing a specific grading agency.

Our econometrics extend and modify the seminal methodology of Moon and Stotsky (1993) in that i) we consider data from three rating agencies (instead of two), and ii) we use a novel formulation of the Monte Carlo Expectation Maximization (MCEM) algorithm (Wei and Tanner, 1990) to circumvent the estimation of multidimensional integrals instead of using probability simulators.

Our results indicate that rating agencies slightly differ in how they weight relevant variables to assess the risk. Most notably, we found a negative and strong correlation

---

at least in part, the informational asymmetry in capital markets, involving insiders possessing more accurate information about the true economic values of their firms (or governments) than outsiders. In turn, rating agencies gain from sharing their information (see Millon and Thakor 1985).

between SNG population (our proxy for number of voters) and debt risk. We interpret this as a *'too-big-to-fail'* situation, i.e., rating agencies consider that large entities are more likely to be bailed out when facing financial problems due to their political power (Hernández et al., 2002). A second strong determinant of ratings is whether the party governing the country is governing the entity under evaluation as well. When the two parties match, debt risk decreases significantly. This is evidence that raters take into account the bailout phenomenon based on both the population of the SNG[6] and political affinity between SNG and federal governments. These results are, to the best of our knowledge, novel in the bond rating literature.

The paper is organized as follows. Section 2 provides a brief description of Mexican intergovernmental relations and reviews the SNG debt environment in Mexico. Section 3 presents a discussion about the opacity of Mexican SNGs. In section 4 we present the model, describe the variables and examine some descriptive statistics. Section 5 discusses the empirical results, and section 6 provides final remarks.

2. A BRIEF OVERVIEW OF MEXICO'S INTERGOVERNMENTAL RELATIONS AND SNG DEBT REGULATION

Mexico is a federal republic composed of three levels of government: the central or federal government, 32 local entities (which consist of 31 states and the Federal District), and 2477 municipalities (SNGs hereafter). Like many countries in the Latin-American region, Mexico is characterized by strong regional and state disparities. While the Federal District and the states of México and Nuevo Leon produce about 40 percent of total GDP,

[6] Population has been interpreted as a political variable in the United States system of federal transfers under the New Deal. For a discussion, see Wallis (1998, 2001) and Fleck (2001).

Chiapas, Guerrero, Hidalgo and Oaxaca reach only a subtotal of 6.8 per cent of total GDP. Clearly, the southern region of the country is by far the poorest.

Mexico follows a *revenue sharing* system where the federal government collects main taxes, namely corporate and personal income taxes, value-added tax, and most excise taxes. These constitute 95 percent of total public sector tax revenue. Through formula, 20 percent of this revenue is redistributed among states and municipalities. These net block transfers are known as *participaciones*. The main deficiencies identified in the system have been the lack of tax independence of local governments and the formula itself. Recently, decentralization efforts have been made. However, this decentralization has not included the revenue side and instead concentrates on expenditures. Moreover, the process has been anarchic and has responded to political pressures and not to efficiency purposes (Hernández 1998).

Regulation of SNG debt is perhaps one of the most important elements to explain its behavior (Ter-Minnasian, 1999). For this reason, we now explain the Mexican case in more detail. First, SNG borrowing is regulated by the national constitution, which specifies that states can only borrow in pesos and solely for productive investment. The details for guaranteeing state credits are contained in the National Fiscal Coordination Law (NFCL), which stipulates that these entities can borrow from commercial and/or development banks and from writing bonds to finance investment projects subject to the previous authorization of the state congress.

Prior to the *tequila* crisis of 1994-95, when a unique political party dominated the country, SNG debt was virtually decided by the federal government in a unilateral manner by direct control of state governments (Díaz 2001). Later, as a consequence of the rapid democratization of the country, this control ended. The new situation allowed states to take advantage of the federal government's concerns for both the banking system (nearly bankrupt as a result of the Tequila crisis), and states' abilities to deliver public services (Hernández, 1998).

Bailouts were common prior to the tequila crisis, though the largest in Mexican history was extended in 1995. As a consequence, virtually no commercial bank developed an institutional capacity to assess sub-national lending. When the *tequila* crisis erupted, most states had high debt ratios and federal bailout occurred.

To correct the situation, the Mexican federal government has faced the challenge to guarantee that bailouts will not occur in the future. This would allegedly be solved by imposing an ex ante market-based mechanism. So a new regulatory framework for debt management by local governments was introduced[7].

States and creditors were induced to make their own trust arrangements in the collateralizing of debt with the block transfers and assuming all the legal risks involved, thus providing recourse for the federal government. A link was established between the risk of bank loans to SNGs and government credit rating.

Currently, credit ratings performed by reputable international agencies are published at global scale. Bank regulators use these ratings to assign capital risk weight for loans provided to states and to municipalities. To control agency shopping, two ratings are mandatory for regulation. In case of a large discrepancy, the capital risk weight of the

[7] Firms (or governments) benefit from obtaining a good rating by lowering the cost of servicing the debt. Many studies for industrial countries have demonstrated empirically that this is generally the case, as they have gained greater acceptance in the market. Ratings have also been used in financial regulation because it simplifies the task of prudential regulation (Cantor and Packer 1995). Thus, as in the Mexican case, regulators have adopted ratings-dependent rules.

lower rate applies. The National Securities Commission recognizes three rating agencies: Moody's, Standard and Poors and Fitch.

The main purpose of the regulation is to discipline SNG debt markets, especially in a new framework characterized by the absence of federal intervention. Financially weaker states and municipalities are likely to be priced or rationed out of the market, while stronger ones would see interest rates on their loans fall (Giugalle et al., 2001).

Another important element in the new regulation is the registration of SNG loans with the federal government. Registration is made conditional upon the borrowing state or municipality being current on the publication of its debt, the related fiscal statistics from the preceding year's final accounts, and on all of its debt service obligations towards the government's development banks. At the same time, in order to make that registration appealing, unregistered loans are automatically risk weighed by the regulators at 150 percent.

Several elements need to be considered to ensure the success of this type of regulation. They include: i) market credibility of the federal commitment about not bailing out defaulting SNGs; ii) quality of the enforcement of capital rules; and iii) quality and reliance of SNG fiscal information, as well as homogeneity in accounting standards.

As we pointed out previously, the largest state in Mexico has been *continuously* bailed out in the past. Furthermore, states and municipalities differ as of today in their accounting standards and not all of them publish their financial statements (Aregional, 2004). These elements cast some doubt on the success of the new regulation.

3. ARE MEXICAN SNGS OPAQUE?

SNG fiscal information is like a black box in Mexico, mainly due to a lack of an adequate institutional and legal framework and a lack of accounting standards[8]. In general, rule of law in Mexico is poor (La Porta and López de Silanes 2001). This problem is larger at state and municipal levels, where transparency is not existent since governments are not required to publish their financial statements (Ugalde 2003).

Transparency issues should be taken into account when rating SNG bonds. Were SNGs transparent, we would not need a lender of last resort since fully transparent states could borrow at market rates that fairly reflected their risk. However, SNG transparency – and thus financial soundness- is more a matter of faith than of fact in Mexico. To discuss this point, we use Morgan's (2002) definition of relative opacity.

This is defined in terms of disagreement between the major bond rating agencies (Fitch, S&P and Moody's) when grading an entity and is used as a proxy for uncertainty. The argument is this: if SNG risk is harder to observe, raters in the business of judging risk should disagree more over SNG bond issues than over other entities. As Table 1 demonstrates, this is the case in Mexican SNGs. This table presents Kappa statistics[9], which are used as a measure of disagreement in biometrics (Cohen 1968). Kappa essentially locates raters along a spectrum between complete disagreement (kappa=0) and complete agreement (kappa=1).

Kappa is 0.13 for the whole set of SNGs (states and municipalities) rated by the three agencies, which suggests a strong disagreement. This figure worsens to 0.05 if only state governments are included. Some SNGs have applied only for two ratings. In this case,

[8] For example, for some municipalities the service of paving roads is registered in current expenditures, whereas for others it is an investment (Hernández, 1998).

[9] Kappa= {po-pe}/{100-pe}, where po is the observed percentage of graded bonds equally; and pe is the expected percentage, given the current distribution of grades.

when the agencies are Fitch and S & P, the kappa is 0.24; if Fitch and Moody's, the figure is 0.17, and finally Moody's and S & P have a 0.04 kappa indicator. These figures suggest that SNGs are opaque in the Morgan (2002) sense. U.S. SNGs rated by Moody's and Fitch have a Kappa 0.61, which suggests that these entities are not as opaque as happens in Mexico

Ederington et al. (1987) suggests that ratings may differ due to three reasons. First, agencies may agree on creditworthiness of a bond but apply different standards for a particular rating. Second, they may differ systematically in the factors they consider and/or the weights attached to each factor. And third, due to the inherent subjectivity of the process, they may give different ratings for random reasons.

In this article, we expect to shed some light on which of the three reasons mentioned above for explaining disparities in ratings is the predominant one when rating sub-national entities in Mexico.

4. EMPIRICAL MODEL AND ESTIMATION

A selectivity problem arises in the analysis of the determinants of SNG bond rating. This follows from the fact that ratings are observed only for those municipalities that have chosen to be rated rather than for all entities in the sample with outstanding debt.

As in Moon and Stotsky (1993) we treat this self-selection problem by developing a model in which we jointly analyze the determinants of the bond rating and the determinants of the decision to obtain a rating. Due to the short history of SNG bond rating in Mexico and in order to gather enough information for our study, we collected ratings from three agencies (Moody's, S&P, and Fitch) instead of the two (Moody's, and S&P) used by Moon and Stotsky. Although estimating a three-agency model is more challenging, it has the advantage of expanding the scope of our conclusions, as we are now able to compare among the rating technologies of more agencies. Additionally, by controlling for tri-variate self-selection, we can consider in the analysis not only SNGs with three ratings but also those with only one or two ratings, and SNGs with no ratings but with outstanding debt.

We also examine jointly the determinants of the bond ratings for the three rating agencies. A joint estimation permits obtaining more efficient estimates by allowing free correlation between selection and rating equations. Allowing free correlation is important since rating decisions are not necessarily independent. Recall that an SNG needs at least two ratings in order to issue a bond registered in the Mexican treasury department, and we are considering three measures of credit risk (the three agencies). Thus, SNG administrators may show preferences for certain agencies if they believe these agencies have a less stringent rating procedure. Additionally, an entity may have incentives to obtain more than one or two ratings if by doing so it lowers the cost of its debt. The literature shows evidence that not only rating values but also their number influence the cost of debt (Ederington et al. 1987). Hence, a multivariate framework applies.

4.1 The model

Following the discussion above, the equation system to solve is:

$$
\begin{aligned}
&y_s^* = X_s\beta_s + \varepsilon_s && \text{propensity to obtain S\&P's rating}\\
&w_s^* = Z_s\gamma_s + \eta_s && S\&P's \text{ perceived riskiness}\\
&y_f^* = X_f\beta_f + \varepsilon_f && \text{propensity to obtain Fitch's rating}\\
&w_f^* = Z_f\gamma_f + \eta_f && Fitch's \text{ perceived riskiness}\\
&y_m^* = X_m\beta_m + \varepsilon_m && \text{propensity to obtain Moody's rating}\\
&w_m^* = Z_m\gamma_m + \eta_m && Moody's \text{ perceived riskiness}
\end{aligned}
\qquad (1)
$$

where index $k = s, f, m$ refers to S&P, Fitch, and Moody's respectively; matrices X_k and Z_k are matrices of explicatory variables; and β_k and γ_k are vectors of parameters to be estimated. The disturbance vector is assumed to be *iid* over entities according to the following six-dimensional normal distribution:

$$
\begin{pmatrix} \varepsilon_{s,i} \\ \eta_{s,i} \\ \varepsilon_{f,i} \\ \eta_{f,i} \\ \varepsilon_{m,i} \\ \eta_{m,i} \end{pmatrix} \sim \mathrm{N}_6\left(\begin{bmatrix} 0\\0\\0\\0\\0\\0 \end{bmatrix}, \begin{bmatrix}
1 & \rho_{\varepsilon_s\eta_s} & \rho_{\varepsilon_s\varepsilon_f} & \rho_{\varepsilon_s\eta_f} & \rho_{\varepsilon_s\varepsilon_m} & \rho_{\varepsilon_s\eta_m} \\
\rho_{\varepsilon_s\eta_s} & 1 & \rho_{\eta_s\varepsilon_f} & \rho_{\eta_s\eta_f} & \rho_{\eta_s\varepsilon_m} & \rho_{\eta_s\eta_m} \\
\rho_{\varepsilon_s\varepsilon_f} & \rho_{\eta_s\varepsilon_f} & 1 & \rho_{\varepsilon_f\eta_f} & \rho_{\varepsilon_f\varepsilon_m} & \rho_{\varepsilon_f\eta_m} \\
\rho_{\varepsilon_s\eta_f} & \rho_{\eta_s\eta_f} & \rho_{\varepsilon_f\eta_f} & 1 & \rho_{\eta_f\varepsilon_m} & \rho_{\eta_f\eta_m} \\
\rho_{\varepsilon_s\varepsilon_m} & \rho_{\eta_s\varepsilon_m} & \rho_{\varepsilon_f\varepsilon_m} & \rho_{\eta_f\varepsilon_m} & 1 & \rho_{\varepsilon_m\eta_m} \\
\rho_{\varepsilon_s\eta_m} & \rho_{\eta_s\eta_m} & \rho_{\varepsilon_f\eta_m} & \rho_{\eta_f\eta_m} & \rho_{\varepsilon_m\eta_m} & 1
\end{bmatrix} \right), \qquad (2)
$$

where $i = 1,...,N$ and N is the sample size. Note that all observations contribute to the estimation of the correlation terms $\rho_{\varepsilon_j\varepsilon_k}$ $j,k = s, f, m$. However, due to self-selection, only those SNGs that have received ratings from the respective agencies contribute to the estimation of terms $\rho_{\varepsilon_j\eta_k}$. Variable $y_{k,i}^*$ is not observable but a binary counterpart, $y_{k,i}$, which takes the value of 1 if $y_{k,i}^* > 0$ and 0 otherwise. The observable counterpart of $w_{k,i}^*$ is categorical ordered so that

$$w_{k,i} = \begin{cases} l_{k,1} & \alpha_{k,1} < w^*_{k,i} \le \alpha_{k,2} \\ l_{k,2} & \text{if} \quad \alpha_{k,2} < w^*_{k,i} \le \alpha_{k,3} \\ \vdots & \vdots \\ l_{k,r} & \alpha_{k,r} < w^*_{k,i} \le \alpha_{k,r+1} \end{cases} , \quad (3)$$

where $l_{k,1} < l_{k,2} \cdots < l_{k,r}$ are consecutive integer values, $\alpha_{k,1} = -\infty$, $\alpha_{k,r+1} = \infty$, and thresholds $\alpha_{k,2} < \alpha_{k,3} < ... < \alpha_{k,r}$ are extra parameters to estimate. In our analysis, we have six categories for all agencies, i.e., $r = 6$ with $l_{k,1} = 0$ and $l_{k,6} = 5 \quad \forall\ k$ (see Table 3).

If $y_{k,i} = 0$, then $w_{k,i}$ does not exist, in accordance with the self-selection mechanism discussed above. Given the binary and categorical ordered nature of the observed counterparts of the dependent variables, parameter identification requires normalization of the diagonal elements in the disturbance covariance matrix as it is presented in (2). Additionally, identification of the coefficients γ_k in the perceived riskiness equations requires either setting to zero one of the thresholds in (3) for each equation or setting the intercept parameter in these equations equal to zero. We chose the first alternative and set $\alpha_{k,2} = 0, \ \ k = s, f, m$.

4.2 Model specification, data and description of variables

In theory, an entity decides to obtain a credit rating because it expects to save enough interest costs to outweigh the agency fee. Thus, level of debt may be a good determinant of the propensity to be rated since the higher the debt, the greater the savings in interests cost (see Moon and Stotsky, 1993).

Likewise, as in most previous literature, we include the total revenue of the entity, as it may represent a good proxy for the local income tax base, and it allows controlling for the size of the entity in terms of economic importance.

If large municipalities in terms of population perceive they will be bailed out, they will then have strong incentives to be rated and obtain debt. We use population as a proxy for size (importance in political terms, after controlling for economic size) since Hernández, Díaz and Gamboa (2002) have shown that populated entities have been bailed out in a more favorable way than unpopulated ones in the past. This variable has also been discussed in the US case. Wallis (1998, 2001) and Fleck (2001) maintain a debate about the political motive of using population on New Deal transfers to states. Since we use a log specification, the inclusion of total revenue and population rules out the possibility of adding revenue per capita as a regressor (to avoid perfect collinearity). Nonetheless, during the estimation process we tried using total revenue and revenue per capita alternatively; we detected neither significant qualitative nor quantitative differences in the final results (except of course in the coefficients of the two regressors).

Finally, we control for political party. We hypothesize that the left wing party has either less financial culture or dismisses market-based approaches with respect to obtaining debt. Thus, dummies for the main political parties were included in the propensity equation.

Regarding the risk assessment equations, the major categories include: (i) political factors; (ii) some indicators of financial soundness including contingent liabilities; (iii) indicators of debt level; and (iv) economic indicators such as gross state product and its composition. Next we describe the variables we consider in our analysis.

Again, population (size in political terms) is a variable that may affect rating behavior. In essence, it may affect it in two ways. First, as previously mentioned, the political decision-making varies with the size of population. Hernández, Díaz and Gamboa (2002) have shown that this variable is a good proxy for the *'too-big-to-fail'* hypothesis when bailing out a state. In this sense, the larger the entity, the higher the number of political votes it has. Second, population is important as a measure of tax base in Mexico. This may be different in advanced economies where smaller municipalities tend to be mostly residential, while larger municipalities tend to have a more substantial industrial base and a more diverse population. In contrast, in LDCs, and Mexico is no exception, small municipalities tend to be more “rural” and thus less subject to being taxed. Complementarily, to control for economic size, we include the entity’s total revenues. This is important since there may be small municipalities in terms of population which are large in terms of economic importance[10].

We include a dummy taking the value one when the political party in control of the municipal government is the same as federal government and taking the value zero otherwise. As already mentioned, we expect that raters allocate a greater probability of bailout to entities with political affinity with the central government and, therefore, it translates to a better risk grade.

For financial soundness, we choose several variables previously used in literature (Ederington, et al., 1987; cantor and Packer, 1996, among others). We include the ratio of an entity’s own revenues to total revenue for two reasons. First, it reflects the flexibility an entity has to absorb a shock; and second, federal transfers to total revenue reflect how compromised the transfer is beforehand. With respect to debt, we use debt to income ratio. Mexican law requires that all new debt must be used in public investment. Thus, one would expect that higher levels of fiscal responsibility imply larger amounts of investment; for this reason, we also include the investment to total expenditure ratio.

---

[10] The typical example in Mexico is San Pedro in the state of Nuevo León.

Regarding the functional form assumed for the model, we follow Moon and Stotsky (1983) and use the log form of all the continuous regressors. The data set contains information from 149 urban municipalities for the year 2001, 148 municipalities for the year 2002, and 147 municipalities for the year 2003. Descriptive statistics of the data are presented in Table 2. We obtain the financial and political variables from INEGI (National Institute of Statistics, Municipal Information System, 2003).

4.3. Estimation approach

It is well known that the main problem in estimating equation systems involving latent variables is the presence of high dimensional integrals in the likelihood function, the highest possible order of integration being equal to the number of latent variables in the system. In contrast to Moon and Stotsky (1993), where the probability simulator of Börsch-Supan and Hajivassiliou (1993) is used, we formulate a Monte Carlo Expectation Maximization (MCEM) algorithm to circumvent the multidimensional integration issue. The main advantages of the MCEM approach are its robustness both to the selection of starting values and to fragile identification (Keane, 1992; Natarajan et al. 2000).

As an introduction to how the MCEM method works, consider the following many-to-one mapping, $\boldsymbol{z} \in Z \to \boldsymbol{y} = y(z) \in Y$. In other words, $\boldsymbol{z}$ is only known to lie in $Z(y)$, and the subset of Z is determined by the equation $\boldsymbol{y} = y(\boldsymbol{z})$, where $\boldsymbol{y}$ is the observed data (variables y_k and w_k in our case), and $\boldsymbol{z}$ is the unobserved information (our y_k^* and w_k^* variables). Thus, the complete data is $\boldsymbol{x} = (\boldsymbol{y}, \boldsymbol{z})$ and the log-likelihood of the observed information is

$$\ell(\theta \mid \boldsymbol{y}) = \ln L(\theta \mid \boldsymbol{y}) = \ln \int_{Z(y)} L(\theta \mid \boldsymbol{x}) d\boldsymbol{z} \quad . \tag{4}$$

Thus, the multidimensional integration problem appears when we try to exclude the unobserved information by integration. Instead of trying to solve (4) directly, the EM algorithm focuses on the complete-information log-likelihood $\ell^c(\theta \mid \boldsymbol{x})$, and maximizes $E\left[\ell^c(\theta \mid \boldsymbol{x})\right]$ by executing two steps iteratively (Dempster et al. 1977). The first one is the so-called Expectation step (E-step), which computes $Q\left(\theta \mid \theta^{(m)}, \boldsymbol{y}\right) = E\left[\ell^c(\theta \mid \boldsymbol{x})\right]$ at iteration $m+1$. The term $E\left[\ell^c(\theta \mid \boldsymbol{x})\right]$ is the expectation of the complete-information log-likelihood conditional on the observed information. provided that the conditional density $f\left(\boldsymbol{x} \mid \boldsymbol{y}, \theta^{(m)}\right)$ is known. The E-step is followed by the Maximization step (M-step), which maximizes $Q\left(\theta \mid \theta^{(m)}, \boldsymbol{y}\right)$ to find $\theta^{(m+1)}$. Then the procedure is repeated until convergence is attained.

The Monte Carlo version of the EM algorithm avoids troublesome computations in the E-step by imputing the unobserved information by Gibbs sampling (Casella and George 1992), conditional on what is observed and on distribution assumptions. In this approach, the term $Q\left(\theta \mid \theta^{(m)}, \boldsymbol{y}\right)$ is approximated by the mean $\frac{1}{K}\sum_{k=1}^{K} Q\left(\theta, \boldsymbol{z}^{(k)} \mid \boldsymbol{y}\right)$, where the $\boldsymbol{z}^{(k)}$ are random samples from $f\left(\boldsymbol{x} \mid \theta^{(m)}, \boldsymbol{y}\right)$. The formulation of an MCEM algorithm for estimating the equation system (1) is presented in Appendix 1.

5. Discussion of empirical results

5.1 Determinants of the rating propensity

Estimation results for the whole set of parameters in the model are given in tables 4a and 4b. We dropped the dummy representing the left-wing political party, PRD, from the regression in order to compare the impact of political orientation on the propensity to be rated. Tables 5 and 6 provide marginal effects of the explanatory variables on propensity-to-be-rated and rating equations, respectively. As is well known, direct discussion of parameter estimates can be misleading in nonlinear models since they measure the impact of the regressors on latent dependent variables, which might have an intuitive meaning but not a definite one (Greene, 2000). Therefore, we focus our discussion on marginal effects, which estimates the effect of regressors on the observed counterparts of the dependent variables for the sample under study. For the particular case of the propensity-to-be-rated equation, the marginal effect provides the change in the probability that an entity requests to be rated as result of a change in the respective regressor. Marginal effects are calculated for each observation; sample averages and standard errors calculated by the delta method are reported.

It turns out that political orientation is important. As observed, the propensity to request a rate increases as we move from the left to the right wing preferences. Thus, it is the PAN (the rightist party) showing the highest propensity. According to Table 5, ceteris paribus, a municipality ruled by the PAN shows a probability to be rated by S&P 16 percentage points (**pp** hereafter) higher than one ruled by the PRD, the leftist party. This figure is approximately 18 for Fitch and decreases to 11 points for Moody's.

This result indicates that entities governed by the PAN are the most willing to obtain a grade. This makes sense in the Mexican case since the PAN is associated with local entrepreneurs, a group with more financial culture (Cabrero, 2004).

Another significant variable that explains propensity to be rated is municipality size measured in population terms. According to Table 5, if municipality A has twice the population of municipality B, then the probability that A asks for the services of S&P

would be about 11 pp[11] higher than B does it. The respective figures for Fitch and Moody's are 10 and 4 pp, all of them significant at the usual levels of significance.

It can be noted that aside from political preferences, population is the most important variable in explaining the decision to be rated. This suggests the (*ex ante*) existence of a self-selection mechanism, where larger municipalities select themselves into the rating process. Regarding financial factors, ratio of own to total revenues is important among those who choose S&P and Fitch, while ratio of debt to total income is important among those who choose Fitch and Moody's.

5.2 Determinants of the Rating

Overall, the estimates support the arguments presented in the motivation of this article. Namely, population, political affinity with the federal government, the ratio of own to total revenues, and the investment variable influence the grade positively (coefficient signs are negative because we assign a lower risk to higher grades; see Table 3). Conversely, the ratio of debt to total income affects the grade negatively.

It can be seen that political variables are important for rating agencies when there is a history of bailout. On the one hand, as we discussed earlier, population size is important because it is politically more costly not to rescue a large entity. On the other hand, the high significance of the dummy for political affinity is evidence that raters allocate a higher bailout probability to entities administrated by the party holding federal office.

For a discussion based on probabilities, Table 6 presents the marginal effects of regressors on the probabilities to receive a given grade (0 to 5 as described in Table 3), conditional on the SNG has requested to be rated.

[11] To get this figure, multiply the corresponding marginal effect by log10 (2) ≈ 0.3.

The regressors that provide statistically significant marginal effects in the S&P rating equation are: political affinity, population, ratio of own to total revenue and ratio of debt to total revenue. Marginal effects for population indicate that a rise in population size shifts the probability distribution from lower to higher grades. In particular, a 10 percent average rise in population brings a 2 pp average increase in the probability to receive an AA or AA^+ grade and 0.4 pp increase in the probability to receive an AA^- from S&P, with simultaneous reductions of 1.0 and 0.8 pp in the probabilities to be rated with A^- or BB, respectively. Although an increase in population favors the probability of obtaining a better grade in Fitch as well, the changes in the distribution of that probability differ from S&P. Thus, a ten percent increase in population implies 0.9 pp reductions in the probabilities of getting A^- and BB and similar 0.6 pp increases in each of the probabilities of obtaining A^+, AA^-, and AA^+. In other words, changes in population tend to have a more uniform impact across the rates in Fitch, while in S&P they tend to affect the lowest and highest rates preferentially. On the other hand, population tested not significant in the Moody's rating equation.

Regarding political affinity, entities governed by the same party as the executive branch have a probability 14 pp higher to get a AA^+, and probabilities 6 pp and 8 pp lower to obtain A^- and BB respectively, than those ruled by a different party. Corresponding numbers are 6 pp, 9 pp, and 10 pp for Fitch, and 3 pp, 14 pp, and 36 pp for Moody's. Again, although the impacts of these determinants may seem to be affecting agencies in a similar qualitative way, they differ quantitatively because they relocate rate probabilities differently across agencies (see Table 6).

5.3 Opacity

We have already demonstrated that agencies seem to take into account the same group of variables when constructing a grade (although Moody's behaves somewhat differently in this regard). However, this condition is not sufficient to ensure that agencies grant an equal grade to a single municipality. Agencies might consider the same factors but they could weight them differently. In what follows we test for SNG opacity by examining whether raters weight the factors in the same way when constructing a grade.

Direct examination of the sample indicates that, among those entities rated by both S&P and Fitch, in only 60% of the events did the two agencies grant the same grade to a particular entity. The proportion is smaller, 44% among those rated by Fitch and Moody's and only 38% among those rated by both S&P and Moody's.

In order to perform a statistical test to detect weighting differences across raters, we compare the parameter estimates (slopes and thresholds) of the rating equations. Three Wald tests comparing the coefficients of the rating agencies by pairs show high statistical differences between S&P and Moody's (chi2-value = 148.1, $p<0.01$) and between Fitch and Moody's (chi2-value = 78.80, $p<0.01$). Smaller but still significant differences were detected between S&P and Fitch (chi2-value = 19. 98, $p<0.05$).

Overall, these results indicate that raters weigh factors in their rating functions differently, which implies there is a high likelihood that they generate different rates even for the same municipality. This is consistent with the kappa analysis presented before.

In the introduction of this article, we mentioned that bond raters have been under scrutiny, especially after the crises in the nineties. Additionally, we noticed that this market has not been subject to study for less developed countries, despite the fact that some doubts about its performance have been expressed (like the Chinese example we provided). Our results suggest that some puzzling grades often observed in LDCs could be explained if we consider not only financial factors in the construction of a rating but also political issues. We have proved that in a country with a history of bailout and opacity, such as Mexico, political variables become important in explaining the grade assigned to the debt of sub national governments.

6. CONCLUSIONS

In this paper, we studied both the determinants of the decision to be rated and the ratings for SNG debt in Mexico, a prominent LDC. One of the main findings is that not only financial but also political factors matter. We showed that population size is a strong determinant of debt rating. In a country with a long bailout history, this result supports our *'too-big-to-fail'* hypothesis. Namely, large entities select themselves to be rated (and so to obtain new debt) because they know they have political power; and secondly, raters know that the probability the federal government will bail out large entities is high. We also showed that political closeness between local and federal governments is important; and rating agencies give a better grade to those entities being governed by the same party as in the executive branch. These outcomes question the purpose of rating sub-national debt in LDCs with a bailout tradition, since the market may assess the risk of these entities as that of sovereign instruments.

Mexico has recently implemented new legislation for the SNG debt market, which pursues increasing the transparency of the market and ruling out debt bailouts. According to our results, it seems that bond rating agencies are not yet convinced of the success of such legislation. Likewise, it is apparent that ratings methodologies take time to evolve and, for the Mexican case at least, they continue echoing the market opacity and bailout tradition of the country.

LITERATURE CITED

Albert, J. and Chib, S. (1993). “Bayesian analysis of binary and polychotomous response data,” *J. of the Am. Stat. Assoc*. 88(42): 669-679

Alfonso A. 2003. Understanding the determinants of sovereing debt ratings: evidence for the two leading agencies. Journal of Economics and Finance 27 (1) 56-74

Aregional. (2004). *Transparencia en estados y Municipios*. México DF.

Bekaert, G. and Campbell, H. (1997). “Emerging equity market volatility.” Journal of Financial Economics 43, 29-77.

Bevilaqua,-Afonso-S; Garcia,-Marcio-G-P (2000). Banks, Domestic Debt, and Crises: The Recent Brazilian Experience Revista-de-Economia-Politica/Brazilian-Journal-of-Political-Economy. Oct.-Dec. 2002; 22(4): 85-103

Bhatia, Ashok-Vir. (2002). “Sovereign Credit Ratings Methodology: An Evaluation.”International Monetary Fund Working Papers: 02/170 2002; 60 pages.

Börsch-Supan, A. and V. Hajivassiliou. (1993). “Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models.” *Journal of Econometrics* 58: 347-368.

Cabrero, Enrique. (2004). *Experiencias en Inovaciones Locales en México.* Editorial Miguel Angel Porrúa, Mexico DF.

Cantor, R. and F. Packer. (1995). “The Credit Rating Industry” *The Journal of Fixed Income* Vol. 5 No. 3.

Cantor, R. and F. Packer. (1996). “Multiple Ratings and Credit Standards: Differences of Opinion in the Credit Rating Industry” *Federal Reserve Bank of New York Staff Reports*. No. 12, NY.

Carleton, W.D. and Lerner, E.M. (1969). “Statistical credit scoring of municipal bonds.” *Journal of Money, Credit and Banking* 1, 4.

Casella, G. and George, E. (1992). “Explaining the Gibbs Sampler.” *The American Statistician* 46 (3): 167-174.

Cohen, J. (1968). “Weighed Kappa: Nominal Scale Agreement with Provision for Scaled Disagreement or Partial Credit.” *Psychological Bulletin* 70 (4).

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). “Maximum Likelihood Estimation from incomplete observations.” *J. Roy. Statist. Soc. B* 39:1-38.

Ederington, L.H., Yawitz J. B., Roberts, B.E. (1987). “The Information Content of Bond Ratings” *Journal of Financial Research* 10 (3).

Fitch. (2002). *“Rating SNGs in México: Methodology”* www.fitchmexico.com/ espanol/Metodologias

Fleck, Robert K. (2001). "Comment: Population, Land, Economic Conditions, and the Allocation of New Deal Spending." *Explorations in Economic History* 38:296-304.

Foster, J., Greer, J., and Thorbecke, E. (1984). “A Class of Decomposable Poverty Measures.” *Econometrica* 52(3).

Giugalle, M., Korobow, A., and Webb, S. (2001). A new Model for Market-Based Regulation of Subnational Borrowing: The Mexican Approach.” Mimeo, The World Bank. Washington DC.

Greene, William H. (2000). *Econometric Analysis*. Prentice Hall. Fourth ed., 1004 pp.

Hernández, F. (1998). “Federalismo Fiscal en Mexico: ¿Como Vamos?” *Revista Internacional de Finanzas Pùblicas*, La Plata, Argentina.

Hernández, F., Díaz, A., and Gamboa, R. (2002). “Bailing out States in Mexico: Causes and Determinants.” *Eastern Economic Journal* 28 (3).

IADB. (1997). *Latin America after a Decade of Reforms*. Research Department Inter-American Development Bank. Washington DC.

Ibrahim, J.G., Chen, M., and Lipsitz, S.R. (2001). “Missing Responses in Generalized Linear Mixed Models when the Missing Data Mechanism is Non-ignorable.” *Biometrika* 88(2): 551-564.

Keane, Michael P. “A Note on Identification in the Multinomial Probit Model”. *Journal of Business and Economic Statistics* 10 (April 1992): 193-200.

Laporta, R. F. López de Silanes, A. Shleiffer and R Vishny. (1999) “Law and Finance” Journal of Political Economy.

Laulajainen, R. (1999). “Subnational Credit Ratings-Penetrating the Cultural Haze.” *GeoJournal* 47, 2.

Levy S. E. Dávila and G. Kessel (2001) “El sur también existe”. *Economía Mexicana,* CIDE, México DF. Vol. VIII.

Louis, T.A. (1982). “Finding the Observed Information Matrix when using the EM Algorithm.” *J. Roy. Statist. Soc. B* 44:226-233.

Meng, X. and Rubin, D. (1993). "Maximum likelihood estimation via the ECM algorithm: a general framework." *Biometrika* 80: 267-278.

Millon, M. and Rubin, D. 1985. "Moral Hazard and Information Sharing: a Model of Financial Information Gathering Agencies. *The Journal of Finance* XL, 5.

Moon, C. G. and Stotsky, G. (1993). "Testing the Differences between the Determinants of Moody's and Standard and Poor's Ratings: An Application of Smooth Simulated Maximum Likelihood Estimation." *Journal of Applied Econometrics* 8, 1.

Morgan, D. (2002) "Rating Banks: Risk and Uncertainty in an Opaque Industry" *American Economic Review.* Volume (Year): 92 Issue (Month): 4 (September)Pages: 874-888

Natarajan, R., McCulloch, Charles E. and Kiefer, N. (2000). "A Monte Carlo EM Method for estimating Multinomial Probit Models." *Computational Statistics and Data Analysis* 34: 33-50.

Reinhart, C. (2001). "Sovereign Credit Ratings Before and After Financial Crises." Mimeo, New York University-University of Maryland Project The Role of Credit Rating Agencies in the International Economy. College Park, MD.

Reinhart, C. (2002). "Default, Currency Crises and Sovereign Credit Ratings." Mimeo, IMF. Washington DC

Remolona E. M. Scatigna and E. Wu. 2008. A Ratings-based approach to measuring sovereign risk. International Journal of Finance and Economics. 13: 26-39.

Sanguinetti, Pablo (2002). Bailing out States in Argentina. Research paper. Interamerican Development Bank. Washington, DC

Serrano (1999 Bailing out Municipalities in Chile. Research paper. Interamerican Development Bank. Washington, DC

Ter-Minnasian, T. (1999). *Fiscal Federalism*. The World Bank Press, Washington DC.

Ugalde, L.C. (2003) "Descentralización y Corrupción en Gobiernos Locales: Una correlacion relevante?" In Descentralización, Federalismo y Planeación Regional en Mexico, edited by R. Tamayo and F. Hérnandez, Porrua Publishing Co. Mexico DF.

Wallis, John Joseph. (1998). "The Political Economy of New Deal Spending, Revisited, With and Without Nevada." *Explorations in Economic History* 35, 140-170.

Wallis, John Joseph. (2001). "The Political Economy of New Deal Spending, Yet Again: A Reply to Fleco." *Explorations in Economic History* 38, 2: 305-314.

Wei, C. and Tanner, M. (1990). "A Monte Carlo implementation of the EM algorithm and the Poor Man's Data Augmentation Algorithms." *J. Am. Statist. Assoc.* 85 (411): 699-704.

Appendix 1.

Let y be a matrix containing all the observed information. The complete information log-likelihood function for the equation system (1) is standard and can be written as the sum of the contributions from eight different regimes. The regimes are represented by: the subsample receiving no grading, the potential three subsamples being graded by a single agency $k = s, f,$ or m, the potential three subsamples being graded by two agencies, and the subsample receiving grades from the all three agencies. The corresponding contributions from the $j = 1,...,8$ regimes to the likelihood are

- regime: $j = 1$: $y_{m,i} = y_{s,i} = y_{f,i} = 0$

$$\ell_j^c\left(\boldsymbol{\theta}_j, \Omega_j \mid \boldsymbol{y}\right) = -\frac{3n_j}{2}\ln\left(2\pi\right) - \frac{n_j}{2}\ln\left|\Omega_j\right| - \frac{1}{2}\mathrm{tr}\left(\Omega_j^{-1}\sum_i \boldsymbol{\varepsilon}_{ji}\boldsymbol{\varepsilon}_{ji}^{'}\right)$$

- regimes $j = 2$: $y_{m,i} = 1;\ y_{s,i} = y_{f,i} = 0$; $j = 3$: $y_{s,i} = 1;\ y_{m,i} = y_{f,i} = 0$; and

$j = 4$: $y_{f,i} = 1;\ y_{m,i} = y_{s,i} = 0$

$$\ell_j^c\left(\boldsymbol{\theta}_j, \Omega_j \mid \boldsymbol{y}\right) = -2n_j\ln\left(2\pi\right) - \frac{n_j}{2}\ln\left|\Omega_j\right| - \frac{1}{2}\mathrm{tr}\left(\Omega_j^{-1}\sum_i \boldsymbol{\varepsilon}_{ji}\boldsymbol{\varepsilon}_{ji}^{'}\right) \quad j = 2,3,4$$

- regimes $j = 5$: $y_{m,i} = y_{s,i} = 1;\ y_{f,i} = 0$; $j = 6$: $y_{m,i} = y_{f,i} = 1;\ y_{s,i} = 0;$; and

$j = 7$: $y_{f,i} = y_{s,i} = 1;\ y_{m,i} = 0$

$$\ell_j^c\left(\boldsymbol{\theta}_j, \Omega_j \mid \boldsymbol{y}\right) = -\frac{5n_j}{2}\ln\left(2\pi\right) - \frac{n_j}{2}\ln\left|\Omega_j\right| - \frac{1}{2}\mathrm{tr}\left(\Omega_j^{-1}\sum_i \boldsymbol{\varepsilon}_{ji}\boldsymbol{\varepsilon}_{ji}^{'}\right) \quad j = 5,6,7$$

- regime $j = 8$: $y_{m,i} = y_{s,i} = y_{f,i} = 1$

$$\ell_j^c\left(\boldsymbol{\theta}_j, \Omega_j \mid \boldsymbol{y}\right) = -3n_j\ln\left(2\pi\right) - \frac{n_j}{2}\ln\left|\Omega_j\right| - \frac{1}{2}\mathrm{tr}\left(\Omega_j^{-1}\sum_i \boldsymbol{\varepsilon}_{ji}\boldsymbol{\varepsilon}_{ji}^{'}\right) \quad j = 8$$

Thus,
$$\ell^c\left(\boldsymbol{\theta}, \Omega \mid \boldsymbol{y}\right) = \sum_{j=1}^{8}\ell_j^c\left(\boldsymbol{\theta}_j, \Omega_j \mid \boldsymbol{y}\right) \tag{5}$$

where $\boldsymbol{\theta} = \left(\beta_m \quad \gamma_m \quad \beta_s \quad \gamma_s \quad \beta_f \quad \gamma_f\right)^{'}$, $\boldsymbol{\theta}_j$ contains the components of $\boldsymbol{\theta}$ present in the equations solved for entities in regime j, Ω_j is the covariance matrix of the disturbance terms associated to those equations j, n_j is the number of observations in regime j, and $\sum_j n_j = N$, the sample size.

E-Step. The expectation of expression (5), conditional on observed information and distribution assumptions, can be written as

$$E\left[\ell^c\left(\boldsymbol{\theta},\Omega \mid \boldsymbol{y}\right)\right]=-\left[\frac{3}{2}n_1+2\left(n_2+n_3+n_4\right)+\frac{5}{2}\left(n_5+n_6+n_7\right)+3n_8\right]\ln\left(2\pi\right)-$$
$$\frac{1}{2}\sum_j n_j \ln\left|\Omega_j\right|-\frac{1}{2}\sum_j \operatorname{tr}\left(\Omega_j^{-1}\sum_i E\left[\boldsymbol{\varepsilon}_{ji}\boldsymbol{\varepsilon}_{ji}^{'}\right]\right)$$

The E-step at iteration $m+1$, requires the calculation of

$$Q_{ji}\left(\boldsymbol{\theta}\mid\boldsymbol{\theta}^{(m)},\Omega_j^{(m)},\boldsymbol{y}\right)=E\left[\boldsymbol{\varepsilon}_{ji}\boldsymbol{\varepsilon}_{ji}^{'}\mid\boldsymbol{\theta}^{(m)},\Omega_j^{(m)},\boldsymbol{y}\right]==\sigma_{ji}^{2(m)}+\begin{pmatrix}\mu_{y_{m,i}^*}^{(m)}-X_{m,i}\beta_m\\ \mu_{w_{m,i}^*}^{(m)}-Z_{m,i}\gamma_m\\ \mu_{y_{s,i}^*}^{(m)}-X_{s,i}\beta_s\\ \mu_{w_{s,i}^*}^{(m)}-Z_{s,i}\gamma_s\\ \mu_{y_{f,i}^*}^{(m)}-X_{f,i}\beta_f\\ \mu_{w_{f,i}^*}^{(m)}-Z_{f,i}\gamma_f\end{pmatrix}\begin{pmatrix}\mu_{y_{m,i}^*}^{(m)}-X_{m,i}\beta_m\\ \mu_{w_{m,i}^*}^{(m)}-Z_{m,i}\gamma_m\\ \mu_{y_{s,i}^*}^{(m)}-X_{s,i}\beta_s\\ \mu_{w_{s,i}^*}^{(m)}-Z_{s,i}\gamma_s\\ \mu_{y_{f,i}^*}^{(m)}-X_{f,i}\beta_f\\ \mu_{w_{f,i}^*}^{(m)}-Z_{f,i}\gamma_f\end{pmatrix}^{'}$$

where $\sigma_{ji}^{2(m)}=\operatorname{Cov}\left(y_{m,i},\ldots,w_{f,i}\mid\boldsymbol{\theta}_j^{(m)},\Omega_j^{(m)},\boldsymbol{y}\right)$, $\mu_{y_{k,i}^*}^{(m)}=E\left[y_{k,i}^*\mid\boldsymbol{\theta}^{(m)},\Omega_k^{(m)},\boldsymbol{y}\right]\quad k=s,f,m$,

$\mu_{w_{k,i}^*}^{(m)}=E\left[w_{k,i}^*\mid\boldsymbol{\theta}^{(m)},\Omega_k^{(m)},\boldsymbol{y}\right]\quad k=s,f,m$. The elements in Q_{ji} associated to equations not solved by entities in regime j must be set equal to zero.

The Gibbs sampler. Gibbs sampling (Casella and George 1992) is necessary to simulate the non-observed information present in the matrices Q_{ji}. The sampler requires the distribution of each $y_{k,i}^*$ and $w_{k,i}^*$ conditional on the values of the rest of the dependent variables in the corresponding regime. It is well known that these distributions are normally univariate under the normality assumption in (2). Let the means and variances of these distributions at the $m+1$ iteration be $\mu_{y_{k,i}^*|\left(-y_{k,i}^*\right)}^{(m)}$, $\sigma_{y_k^*|\left(-y_k^*\right)}^{2(m)}$, $\mu_{w_{k,i}^*|\left(-w_{k,i}^*\right)}^{(m)}$, and $\sigma_{w_k^*|\left(-w_k^*\right)}^{2(m)}$, respectively, where $|\left(-y_{k,i}^*\right)$ indicates conditionality on the values of all the other dependent variables (apart from $y_{k,i}^*$) being present in the regime at which entity i belongs.

Simulations for $y_{k,i}^*$ must be done conditional on its corresponding observed information $y_{k,i}$. The observed counterpart of $y_{k,i}^*$ is dichotomous with $y_{k,i}^*$ being positive if $y_{k,i}$ equals one and non-positive if

$y_{k,i}$ equals zero. Accordingly, we simulate $y^*_{k,i}$ from a normal distribution with mean $\mu^{(m)}_{y^*_{k,i}|(-y^*_{k,i})}$ and variance $\sigma^{2(m)}_{y^*_k|(-y^*_k)}$ truncated below at zero if $y_{k,i}$ equals one and truncated above at zero if $y_{k,i}$ equals zero. The observed counterparts of variables $w^*_{k,i}$ are categorical ordered and defined by (3). Correspondingly, we simulate $w^*_{k,i}$ from a normal distribution with mean $\mu^{(m)}_{w^*_{k,i}|(-w^*_{k,i})}$, and variance $\sigma^{2(m)}_{w^*_k|(-w^*_k)}$ truncated above at $\alpha_{k,t+1}$ and truncated below at $\alpha_{k,t}$ when $w_{k,i}$ equals $l_{k,t}$ ($k = s, f, m;\ t = 1,...,r$).

A complete set of starting values $y^{*(0)}_{k,i}$ and $w^{*(0)}_{k,i}$ is required to initiate the Gibbs sampler. We use $y^{*(0)}_{k,i} = 0 \ \ \forall k, i$ and $w^{*(0)}_{k,i} = w_{k,i}$. The simulation was then repeated iteratively until completing sequences $y^{*(1)}_{k,i},..., y^{*(K(m))}_{k,i}$ and $w^{*(1)}_{k,i},..., w^{*(K(m))}_{k,i}$, where $K^{(m)}$ is a number large enough to ensure convergence. Wei and Tanner (1990) recommend starting with a small $K^{(1)}$ and progressively increasing $K^{(m)}$ as m increases. Then, eliminate a number k_{burn} of simulations from the beginning of the sequence. The remaining simulations in the sequence are used to estimate the terms $\sigma^{2(m)}_{ji}$, $\mu^{(m)}_{y^*_{k,i}}$, and $\mu^{(m)}_{w^*_{k,i}}$ in Q_{ji}.

M-Step. Following Meng and Rubin (1993), it is advisable to replace the M-step by two conditional M-steps. The first conditional M-step maximizes $E\left[\ell^c\left(\boldsymbol{\theta}, \Omega \mid \boldsymbol{y}\right)\right]$ with respect to the elements in $\boldsymbol{\theta}$ conditional on $\boldsymbol{\theta}^{(m)}$ and $\Omega^{(m)}$. After a little matrix calculus, it is easy to see that the maximizer in this first conditional maximization can be written as a generalized least squares estimator

$$\boldsymbol{\theta}^{(m+1)} = \left[X_d^{'} \left[\sum_j \left(\tilde{\Omega}_j^{-1} \otimes I^j \right) \right] X_d \right]^{-1} X_d^{'} \left[\sum_j \left(\tilde{\Omega}_j^{-1} \otimes I^j \right) \right] \mu^{(m)}_{y^*}$$

where I^j is a $N \times N$ diagonal matrix with $I^j_{ii} = 1$ if entity i belongs to regime j and $I^j_{ii} = 0$ otherwise. The 6x6 matrix $\tilde{\Omega}_j^{-1}$ contains the elements of Ω_j^{-1} in the positions corresponding to the equations solved in regime j, while the remaining elements must be set equal to zero. The block-diagonal matrix X_d is defined as

$$X_d = \begin{bmatrix} X_m & 0 & \cdots & 0 \\ 0 & Z_m & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & Z_f \end{bmatrix}$$, where $Z_{k,i} = 0$ if $y_{k,i} = 0$; $\mu_{y^*}^{(m)} = \left(\mu_{y_m^*}^{(m)} \quad \mu_{z_m^*}^{(m)} \quad \cdots \quad \mu_{z_f^*}^{(m)}\right)^T$,

$\mu_{y_k^*}^{(m)} = \left(\mu_{y_{k,1}^*}^{(m)} \quad \cdots \quad \mu_{y_{k,i}^*}^{(m)} \quad \cdots \quad \mu_{y_{k,N}^*}^{(m)}\right)^T$, $\mu_{z_k^*}^{(m)} = \left(\mu_{z_{k,1}^*}^{(m)} \quad \cdots \quad \mu_{z_{k,i}^*}^{(m)} \quad \cdots \quad \mu_{z_{k,N}^*}^{(m)}\right)^T$ and $\mu_{z_{k,i}^*}^{(m)} = 0$ if $y_{k,i} = 0$

The second conditional M-step estimates $\Omega^{(m+1)}$ by maximizing $E\left[\ell^c\left(\boldsymbol{\theta}, \Omega \mid \boldsymbol{y}\right)\right]$ with respect to the elements in Ω conditional on $\boldsymbol{\theta}^{(m+1)}$ and $\Omega^{(m)}$. No closed form for $\Omega^{(m+1)}$ exists; thus, numerical optimization techniques must be used at this stage. Thresholds $\alpha_{k,3} < \ldots < \alpha_{k,r}$ are not present in the complete-information likelihood function; therefore, they cannot be obtained by first order condition or by numerical optimization. We proceed the following way to estimate $\alpha_{k,t}$: i) at every round of the Gibbs sampler at iteration m, keep the minimum value of every sequence obtained when simulating the observations $w_{k,i} = l_{k,t}$; this produces a set of $K^{(m)} - k_{burn}$ values; ii) keep the maximum value of every sequence obtained when simulating the observations $w_{k,i} = l_{k,t-1}$; iii) take the medians of the two sets obtained in (1) and (2); iv) take the average between the two medians, which produces a consistent estimator of $\alpha_{k,t}$. The E-step and M-step are then repeated until convergence is attained.

Appendix 2. The Information matrix

Louis's identity (Louis 1982) was used in this study to obtain a Monte Carlo estimation of the information matrix

$$I\left(\boldsymbol{\theta};\boldsymbol{y}\right) = -H^c\left(\boldsymbol{\theta};\boldsymbol{x}\right) - E\left[S^c\left(\boldsymbol{\theta};\boldsymbol{x}\right)S^c\left(\boldsymbol{\theta};\boldsymbol{x}\right)'\right] + E\left[S^c\left(\boldsymbol{\theta};\boldsymbol{x}\right)\right]E\left[S^c\left(\boldsymbol{\theta};\boldsymbol{x}\right)'\right]$$

where $H^c\left(\boldsymbol{\theta};\boldsymbol{x}\right) = \dfrac{\partial^2 \ell^c\left(\boldsymbol{\theta};\boldsymbol{x}\right)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}$ and $S^c\left(\boldsymbol{\theta};\boldsymbol{x}\right) = \dfrac{\partial \ell^c\left(\boldsymbol{\theta};\boldsymbol{x}\right)}{\partial\boldsymbol{\theta}}$ are the complete information Hessian and Score vector, respectively. All of the expectations are estimated at the final MCEM estimators. Monte Carlo estimates of the complete information Hessian and score can be used to estimate the information matrix (Ibrahim et al. 2001).

Since thresholds $\alpha_{k,t}$ are not present in the complete-information maximum likelihood, their standard errors cannot be obtained from the information matrix presented above. Following Albert and Chib (1993), we consider that estimates of $\alpha_{k,t}$ are uniformly distributed between the two medians calculated in step (3) when estimating $\alpha_{k,t}$ in Appendix 1. Thus, standard errors for our estimates of $\alpha_{k,t}$ were calculated as the square roots of the variances of those distributions.

Table 1. Kappa index

USA: (Morgan, 2002)	Mexico:
Banks = 0.30 Other Sectors = 0.45	Banks = 0.27 Other sectors = 0.36 States and Municipalities = 0.13

Table 2. Descriptive statistics of dependent and explanatory variables

Binary dependent variables		Sum	
S&P	The entity was rated by S&P in the period 2001-2003 (yes=1)	96	
Fitch	The entity was rated by Fitch in the period 2001-2003 (yes=1)	74	
Moody's	The entity was rated by Moody's in the period 2001-2003 (yes=1)	40	
Dummy explanatory variables		**Sum**	
PRD	The entity is administered by the PRD party (yes=1)	45	
PRI	The entity is administered by the PRI party (yes=1)	148	
COA	The entity is administered by a COALITION party (yes=1)	60	
PAN	The entity is administered by the PAN party (yes=1)	191	
A	The entity is administered by the same party as federal government (yes=1)	191	
Continuous explanatory variables[1]		**Mean**	**Std. dev**
Pop	2000 Population ($x10^5$)	3.3	3.1
TI	Total annual income ($US\$x10^8$)	25.7	31.7
O_T	Own to total revenue ratio	0.23	0.12
D_I	Debt to income ratio	0.12	0.17
Debt	Total debt ($US\$x10^6$)	20.8	44.8
P_D	Per capita debt ($US\$x10^3$)	0.58	0.90
I_G	Investment to total expenditure ratio	0.23	0.13

[1] The $\log_{10}$ form of the continuous explanatory variables was used in the estimation.

Table 3. Equivalence between ordinal and qualitative rates

	Rating Institution		
Ordinal rate	**S&P**	**Fitch**	**Moody's**
0	AA+, AA	AA	Aa2
1	AA-	AA-	Aa3
2	A+	A+	A1
3	A	A	A2
4	A-	A-,A3	A3
5	BB+,BB-	BBB+,BBB	Baa1,Bba1

Table 4.a. Model estimates

		S&P		Fitch		Moody's	
Equation	**Variable**	**Estimate**	**Std. error**	**Estimate**	**Std. error**	**Estimate**	**Std. error**
Propensity to be rated	**Constant**	-1.4967 [a]	0.3798	-1.0241 [a]	0.3944	-4.9557 [a]	0.4660
	PRI	0.3496	0.2497	0.4515 [b]	0.2033	3.2767 [a]	0.4196
	Coalition	0.5976 [b]	0.2941	0.3982	0.2815	3.4165 [a]	0.4330
	PAN	0.8437 [a]	0.2611	0.9766 [a]	0.2011	3.4630 [a]	0.4176
	POP	1.8514 [a]	0.3395	1.6517 [a]	0.3202	1.0679 [b]	0.4313
	TI	-0.0251	0.2197	-0.4905 [b]	0.2361	0.0638	0.2566
	O_T	1.1772 [a]	0.2541	0.7546 [a]	0.2413	-0.0828	0.2801
	D_I	-0.0072	0.0573	0.1878 [a]	0.0512	0.2601 [a]	0.0633
Rating	**Constant**	0.9318	0.8961	0.8772	1.0615	1.9660	1.3718
	A	-0.7958 [a]	0.2494	-0.7496 [b]	0.3714	-1.9033 [a]	0.3383
	Pop	-2.3515 [a]	0.6811	-1.7098 [c]	0.9723	-1.5152	1.8990
	TI	0.2692	0.3831	0.1023	0.5397	0.4503	1.0746
	O_T	-2.5806 [a]	0.8021	-2.8562 [a]	0.8788	-2.3790 [a]	0.9092
	D_I	0.1255	0.0925	0.0378	0.1409	0.1500	0.1873
	I_G	-1.0191 [b]	0.4320	-1.2384 [b]	0.5175	-1.5604 [b]	0.6599

Test for model significance (restricted model: slopes are all zero) Chi2-value = 710.4315 gl=39 ($p<0.01$)

[a] significant at 1% significance; [b] significant at 5% significance; [c] significant at 10% significance.■

Table 4.b. Thresholds and covariance matrix

		S&P		Fitch		Moody's	
		Estimate	**Std. error**	**Estimate**	**Std. error**	**Estimate**	**Std. error**
Thresholds	$\alpha_{k,3}$	0.5030 [a]	0.1242	0.8303 [a]	0.1732	2.6975 [a]	0.3094
	$\alpha_{k,4}$	1.4025 [a]	0.1205	1.8727 [a]	0.1200	3.0433 [a]	0.2239
	$\alpha_{k,5}$	2.0564 [a]	0.1146	2.4665 [a]	0.1385	3.8627 [a]	0.3222
	$\alpha_{k,6}$	2.9212 [a]	0.2337	3.5228 [a]	0.2850	4.5533 [a]	0.2924

		Estimate	Std. error		Estimate	Std. error
Covariance matrix	$\rho_{\varepsilon_s \eta_s}$	-0.2675	0.1898	$\rho_{\eta_s \eta_m}$	0.7370 [a]	0.1229
	$\rho_{\varepsilon_s \varepsilon_f}$	0.6536 [a]	0.0387	$\rho_{\varepsilon_f \eta_f}$	-0.1022	0.5946
	$\rho_{\varepsilon_s \eta_f}$	-0.0745	0.3518	$\rho_{\varepsilon_f \varepsilon_m}$	0.0442	0.0778
	$\rho_{\varepsilon_s \varepsilon_m}$	0.2840 [a]	0.0635	$\rho_{\varepsilon_f \eta_m}$	-0.1963	0.4274
	$\rho_{\varepsilon_s \eta_m}$	-0.4171	0.4097	$\rho_{\eta_f \varepsilon_m}$	-0.0662	0.1986
	$\rho_{\eta_s \varepsilon_f}$	0.1288	0.2232	$\rho_{\eta_f \eta_m}$	0.6897 [a]	0.2519
	$\rho_{\eta_s \eta_f}$	0.6351 [a]	0.1291	$\rho_{\varepsilon_m \eta_m}$	0.0218	0.3303
	$\rho_{\varepsilon_m \eta_m}$	-0.1191	0.1832			

[a] significant at 1% significance; [b] significant at 5% significance; [c] significant at 10% significance.■

Table 5. Marginal effects for the propensity-to-be-rated equations

	S&P		Fitch		Moody's	
Variable	**Estimate**	**Std. error**	**Estimate**	**Std. error**	**Estimate**	**Std. error**
PRI	0.0553	0.0354	0.0611 [b]	0.0249	0.0815 [a]	0.0164
Coalition	0.1052 [b]	0.0482	0.0521	0.0387	0.1021 [a]	0.0288
PAN	0.1633 [a]	0.0400	0.1771 [a]	0.0290	0.1097 [a]	0.0168
POP	0.3873 [a]	0.0703	0.3378 [a]	0.0658	0.1498 [b]	0.0633
TI	-0.0053	0.0460	-0.1003 [b]	0.0485	0.0090	0.0360
O_T	0.2462	0.0542	0.1543 [a]	0.0508	-0.0116	0.0393
D_I	-0.0015 [a]	0.0120	0.0384 [a]	0.0107	0.0365 [a]	0.0094

[a] significant at 1% significance; [b] significant at 5% significance; [c] significant at 10% significance.■

Table 6. Marginal effects for the rating equations

Variable	Rate	S&P		Fitch		Moody's	
		Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
A	**0**	0.1471 [a]	0.0472	0.0588 [b]	0.0275	0.0307 [c]	0.0175
	1	0.0439 [b]	0.0210	0.0797 [b]	0.0371	0.4918 [b]	0.1939
	2	0.0091	0.0204	0.0747 [c]	0.0429	0.0338	0.0606
	3	-0.0534 [b]	0.0226	-0.0144	0.0134	-0.0487	0.0742
	4	-0.0845 [b]	0.0394	-0.0966 [b]	0.0482	-0.1441	0.1033
	5	-0.0621 [c]	0.0336	-0.1022	0.0639	-0.3634	0.2289
Pop	**0**	0.4870 [a]	0.1426	0.1682 [b]	0.0786	0.0608	0.0906
	1	0.1102 [a]	0.0343	0.1893 [b]	0.0847	0.2831	0.3027
	2	-0.0096	0.0477	0.1349 [c]	0.0719	-0.0007	0.0149
	3	-0.1622 [a]	0.0572	-0.0573	0.0391	-0.0346	0.0547
	4	-0.2401 [a]	0.0871	-0.2247 [b]	0.1059	-0.0587	0.0731
	5	-0.1853 [b]	0.0855	-0.2104 [c]	0.1108	-0.2499	0.2818
TI	**0**	-0.0627	0.0904	-0.0072	0.0470	-0.0178	0.0456
	1	-0.0149	0.0216	-0.0079	0.0530	-0.0829	0.1928
	2	0.0000	0.0061	-0.0053	0.0374	0.0002	0.0044
	3	0.0209	0.0302	0.0026	0.0165	0.0102	0.0248
	4	0.0318	0.0461	0.0094	0.0627	0.0172	0.0393
	5	0.0250	0.0372	0.0084	0.0583	0.0732	0.1749
O_T	**0**	0.5662 [a]	0.1826	0.2955 [a]	0.0991	0.0943	0.0689
	1	0.1315 [a]	0.0483	0.3337 [a]	0.1175	0.4389 [a]	0.1583
	2	-0.0055	0.0547	0.2398 [b]	0.1102	-0.0011	0.0224
	3	-0.1885 [b]	0.0750	-0.0999	0.0655	-0.0537	0.0542
	4	-0.2831 [a]	0.1098	-0.3961 [a]	0.1494	-0.0911	0.0756
	5	-0.2207 [b]	0.1051	-0.3730 [b]	0.1482	-0.3873 [a]	0.1488
D_I	**0**	-0.0294	0.0216	-0.0053	0.0126	-0.0058	0.0093
	1	-0.0070	0.0052	-0.0061	0.0143	-0.0269	0.0343
	2	0.0000	0.0028	-0.0046	0.0106	0.0001	0.0014
	3	0.0098	0.0074	0.0017	0.0044	0.0033	0.0063
	4	0.0149	0.0115	0.0073	0.0170	0.0056	0.0078
	5	0.0118	0.0097	0.0071	0.0162	0.0237	0.0299
I_G	**0**	0.2409 [b]	0.0987	0.1305 [b]	0.0620	0.0619	0.0386
	1	0.0577 [b]	0.0281	0.1476 [b]	0.0650	0.2880 [b]	0.1439
	2	0.0005	0.0232	0.1063 [c]	0.0610	-0.0007	0.0145
	3	-0.0801 [b]	0.0390	-0.0440	0.0285	-0.0353	0.0311
	4	-0.1224 [c]	0.0626	-0.1751 [b]	0.0867	-0.0598	0.0565
	5	-0.0966 [c]	0.0526	-0.1652 [c]	0.0866	-0.2542 [c]	0.1317

[a] significant at 1% significance; [b] significant at 5% significance; [c] significant at 10% significance.■