Evidence for Relational Contracts in Sovereign Bank Lending *

Peter Benczur[†] Cosmin L. Ilut[‡]

October 2009

Abstract

This paper presents direct evidence for relational (self-enforcing dynamic) contracts in sovereign bank lending. Unlike the existing empirical literature, its instrumental variables method allows for distinguishing a direct influence of past repayment problems on current spreads (a "punishment" effect in prices) from an indirect effect through higher expected future default probabilities. Such a punishment provides positive surplus to lenders after a default, a feature that characterizes relational contracts. Using developing country data for 1973-1981 and constructing continuous variables for credit history, we find evidence that most of the influence of past repayment problems is through the direct, punishment channel.

Key Words: reputation, relational contracts, sovereign bank loan spreads, rational expectations, default risk.

JEL Classification: C73, D86, F34, G12, G14, G15.

---

*We thank Craig Burnside, Larry Christiano, Gábor Kézdi, Narayana Kocherlakota, Sergio Rebelo, Ádám Szeidl, Balázs Szentes, Mark Wright and seminar participants at the National Bank of Hungary (MNB) and Northwestern University for helpful comments and suggestions. Part of the work was done while Ilut was a visiting scholar at the National Bank of Hungary (MNB).

[†]National Bank of Hungary (MNB) and Central European University, email: benczurp@ceu.hu

[‡]Duke University. Corresponding author, e-mail: cosmin.ilut@duke.edu

1 Introduction

A central issue in limited contract enforceability is that if the identity of a defaulting party is forgotten, then it can reenter the market with a fresh start. The acts of countries and major banks, however, are remembered, thus they cannot simply walk away after breaking their commitments. This implies that one can view sovereign bank lending as a long-term relationship (to use the terminology of Baker, Gibbons and Murphy, 2002, a relational contract) between borrowers and lenders. In such contracts, parties honor their obligations in order to influence the terms of future interactions. This can mean that a default or any other form of misbehavior leads to an increase in borrowing costs or, as an extreme form, a capital market exclusion. As shown by Kletzer and Wright (2000), Wright (2002), and Kovrijnyikh and Szentes (2007), such punishments are fully compatible with competitive markets, due to the repeated lender interactions that characterize relational contracts..

There is indeed evidence that a default leads to an increase in spreads (Ozler, 1993, Eichengreen and Mody, 1999, Reinhart et al, 2003). This literature, however, has been unable to identify the precise mechanism of this effect. In particular, it cannot tell apart a relational contract argument from a signaling alternative: a default reveals some adverse information about the expectation of the debtor's future output (for example, the type of the debtor), which hurts its future outcomes (Eaton, 1996, Sandleris, 2008). This mechanism can be labeled as "domestic costs", or "reputational spillovers" in the classification of Panizza et al (2009).

Can we separate the two mechanisms? Our identification strategy is based on the following observation. In the signaling case, reputational spillovers indirectly contribute to a higher spread by increasing the probability of a future default, but lenders still earn zero expected profits. In contrast, in relational contracts punishments are incorporated directly into future prices, giving positive surpluses to lenders. The existing empirical literature cannot determine whether a default is followed by an increase in sovereign bank spreads in excess of the increase of future nonrepayment risk. We present empirical evidence that a sovereign default is followed by positive lender surpluses, which is consistent with the relational contract mechanism.

Such evidence has immediate consequences for understanding sovereign risk, as it points to the presence of dynamic incentives as a repayment mechanism. It is also relevant for the broad context of repeated games and reputations: besides prices and markets, "relationships provide an alternative mechanism that also plays an important role in allocating resources "(Samuelson, 2006, Section 1.3). In both cases, there is very little direct evidence on dynamic incentives themselves - the sovereign risk literature, for example, usually calibrates its models

to match various aggregate outcomes, like the cyclicality of sovereign debt flows and interest rate spreads, or the timing and frequency of default.

The purpose of our paper is to contribute to the empirical sovereign risk literature in two major respects. First, and most importantly, we focus on the distinction of any direct effect of a bad repayment history on the spread, above the indirect one going through increased default probability. Second, we deploy different econometric techniques and variables in order to control for country fixed effects, a problem that is important in the framework of this analysis and has not been thoroughly dealt with in most of the previous research.

The data on spreads is from the World Bank's publication "Borrowing in International Capital Markets" for the period 1973-1981, on 37 developing countries. This period was the heyday of syndicated bank lending to sovereigns. While most of the variables utilized in the paper are those suggested by the literature, we create continuous measures of past and future default, which are based on arrears data from the World Bank's Global Development Finance. These variables are compatible with country fixed effects.

The main estimation strategy used in the paper is a structural asset pricing approach. The starting point is that the spread is determined by expected default risk and credit history. As the former is an unobserved variable, we replace the expectation term with its realization and a prediction error. This creates an identification problem, as the realization is correlated with the prediction error. Assuming that expectations are rational, we can use any variable (debtor characteristics) available at the time of pricing as a valid instrument.

This procedure basically attributes the reduced form explanatory power of certain fundamentals to their predictive power for future default. Overidentification thus receives a central role: it tests whether information used for pricing affects the spread only through predicted default or there is an extra channel of influence. In particular, we first show that if we consider default risk as the only determinant of the spread, the p value of the overidentification test is small. Our approach is that the rejection of the overidentification test indicates that the instruments are proxying for a missing term in the structural equation, which is due to the direct effect of some factor on the spread, above the one coming from the influence on the predicted probability. Indeed, when we add credit history as an extra right hand side variable, we find that the p value of the overidentification test increases substantially, and that the coefficient on credit history is positive and significant. This means that after a default, the spread increases more than justified by increased future risk, indicating a positive surplus for lenders. Finally, we search for additional variables which are pivotal for overidentification.

In the reduced form estimation, we find that both recent and distant default history have a significant positive influence on the spread, but the inclusion of country fixed effects is

necessary. The conclusion of our benchmark structural specification is that future default risk, recent default history and an extra term given by reserves to imports, can robustly and meaningfully describe the spread.

In the structural form, we find that the coefficient estimate for default risk is very similar to the one in the reduced form. Its magnitude is around 0.35. This means that if we increase our default risk measure from its median to the 90th percentile, the spread goes up by 17 basis points. Compared to the average spread of 133 basis points, it can be considered as sizable. Similarly, raising the reserves to imports ratio from its median to the 90th percentile, the spread goes up by 24 basis points. For credit history to affect the spread in a quantitatively meaningful way, we need to look at countries with an extremely poor record: for certain countries in our sample, there is a 32-56 basis point direct punishment component in the spread (keeping default risk unchanged). In an illustrative example based on Brazilian numbers, the present value of the reward for clearing all arrears is quite comparable to the amount of arrears themselves. The structural form estimation offers strong statistical and economic evidence that credit history has a dominantly direct effect on loan spreads.

The structure of the paper is the following: the second section comprises a review of the theoretical and empirical literature on the role of credit history. The next part explains the empirical strategy of the structural asset-pricing setup. The description of data, variables and the main econometric problems is presented in the fourth section. The fifth part describes the reduced- and structural-form results, while the last section concludes.

2 Related theoretical and empirical literature

The sovereign risk literature has identified many channels through which a default can influence future borrowing terms. One interpretation is based on signaling: a default reveals some information about the debtor, which hurts its future outcomes (borrowing terms or third party decisions like private investment), so the debtor avoids default in order to send a favorable signal about its fundamentals (as in Sandleris, 2008) or type (as in Eaton, 1996).

Alternatively, one can view sovereign borrowing as a relational contract between borrowers and lenders. In this case, there is an implicit or explicit agreement on lending and repayment terms, and any deviation (default) would initiate some punishment. Reputation as a repayment incentive then refers to the case of indirect punishments, like exclusion from future borrowing. In most cases (like Eaton and Gersovitz, 1981, Kletzer and Wright, 2000, Wright, 2002, Yue, 2006), this punishment is an out-of-equilibrium threat. In some models this is already sufficient to ensure full repayment (Eaton and Gersovitz, 1981, Kletzer and Wright, 2000, Wright, 2002). Consequently, there is no default or punishment in equilibrium.

In Yue (2006), the punishments are the threat points of the renegotiation process, so there is default in equilibrium, but no punishment. Note that being an out-of-equilibrium threat need not imply the lack of credibility: both in Kletzer and Wright (2000) and Wright (2002), the threats are subgame perfect and renegotiation proof.

There are some models with default and/or punishment along the equilibrium path. In most cases, however, default is a consequence of incomplete markets (debt contracts are non-contingent). Examples include Yue (2006), Sandleris (2008), and Arellano (2008). The recent literature on unsecured household debt (for example, Chatterjee et al, 2007) also allows for default, with an exogenously set bankruptcy procedure (unrelated to reputation) as the punishment. There is equilibrium default in the investment model of Hopenhayn and Werning (2008), but there is no direct punishment after a default, the relationship simply ends.

Relaxing the assumption of perfect observability of parties' actions (like bank fees kept confidential), the theory of repeated games with imperfect monitoring implies that even contingent contracts could feature occasional episodes of punishment and potentially "default" as well: in the cartel literature, Green and Porter (1984), and Tedeschi (1994) establish theoretically that collusive and noncooperative behavior is mixed, even in the absence of actual cheating (default) in the cartel.[1] Adding some uncertainty about the true realization of output in the Kletzer and Wright (2000) model is likely to lead to default and punishment along the equilibrium path.

In Kovrijnyikh and Szentes (2007), there is no explicit breaking of the contract. Instead, they have a "debt hangover" situation, when the borrower cannot fully repay in a given period, which gives a de facto monopoly power to the incumbent lender. After "writing off" part or all of the initial debt, the incumbent can neglect other lenders and can extract some monopoly rents ("endogenous exclusion from competitive markets"). If one interprets a debt hangover as a full write-off of the initial debt (which is a sunk cost for the incumbent), then the monopoly contract in fact implies an interest rate giving extra profits to the incumbent.

It is not obvious that a competitive capital market is compatible with nonzero lender surpluses after a default, as a competing lender might be tempted to offer a cheaper loan. In dynamic relations, however, positive surpluses can be maintained by repeated lender interactions. In Kletzer and Wright (2000), the punishment is compatible with external competition, due to a "cheat the cheater" response of the other lenders. This leads to an implicit seniority of preexisting loans. Kovrijnyikh and Szentes (2007) assume such a seniority explicitly, which gives the incumbent lender monopoly power after a debt hangover. The common feature is that preexisting debt limits the impact of outside competition, and

[1] Abreu, Pearce and Stacchetti (1990) generalize these results to discounted repeated games.

allows for punishments which give positive surplus to the lender even in the case of potential entrants. An alternative mechanism is described by Wright (2002): for syndicated loans, in which each bank has a share of the profits, then each bank's incentive to maintain a good reputation in this cooperation makes them tacitly collude in punishing a country in default.

Switching to empirics, there is a diverse literature aimed at detecting behavior in line with dynamic incentives. In the economic history literature, Greif (1993) looks at the coalition of Maghribi traders in medieval times, and finds that their merchant-agent relationships evolved consistently with relational contracts (a repeated game with imperfect monitoring). Milgrom, North and Weingast (1990) also point to the role of self-enforcing contracts in medieval trade. Porter (1983) finds empirical support for repeated games with imperfect monitoring (switches between collusive and noncollusive behavior in the US railroad industry), while Levenstein (1997) performs a similar analysis for the bromine industry in the US.

Focusing on sovereign borrowing, there is some direct evidence on the repayment incentives of a sovereign debtor. This literature is surveyed and discussed in Panizza et al (2009). The more relevant issue for our discussion is how a country's credit history affects the borrowing cost of a sovereign. Eichengreen and Mody (1999) use data on 4500 loans over the 1991-1997 period and employ a pooled OLS regression, corrected for sample selectivity. They notice that a history of debt reschedulings has a weak positive effect on the probability of an issue while it significantly increases the spread that successful issuers pay.

Ozler (1993) is an important contribution to the issue of reputation, which has been cited by many theoretical papers as the main evidence of an effect of repayment history on credit terms.[2] She uses data on 64 countries for the period 1968-1981, which was one of rapid international lending expansion. The econometric technique is a pooled OLS regression with time-specific dummies. Her main results are that the spread is influenced by relatively recent repayment history (which she identifies as the 1930s through 1960s), but not by distant history (before 1930).

Reinhart et al (2003) provides a documentation of the effect of repayment behavior on sovereign debt. Employing a cross-sectional regression with multiyear averages of measures for default risk, history of repayment, inflation rates and external debt as controls, they find that a history of defaults weakens a country's ability to borrow large amounts on reasonable terms, because a bad credit history is reflected in lower creditworthiness (proxied by the country's credit rating).

Our analysis extends that of Ozler (1993) and Reinhart et al (2003) in two major ways. The first one concerns the treatment of country fixed effects. By including time invariant variables like dummies for repayment problems, Ozler can no longer have country fixed

[2]See for example Eaton (1996) or Obstfeld and Rogoff (1996, p.379).

effects. We resolve this issue by constructing a continuous measure of recent default, which allows us to use both country fixed effects and default history indicators.

The second, more important contribution is a structural and causal empirical approach. The ability to distinguish between different channels of influence is particularly important for the credit history case. As we argued before, there are two ways in which history could affect the spread. Looking only at the reduced-form results – as Ozler (1993) and Reinhart et al (2003) do – one cannot separate the two effects.

3 The empirical strategy: identification in a structural asset pricing regression

The starting point is that the spread reflects perceived risks and potential extra effects (like a punishment surcharge for past defaults). Such a specification thus involves latent expectations of the risk(s) based on information at the time of pricing. Three main solutions have been adopted to overcome this issue. One widely used approach has been to assume specific functional form relations between the spread, the risk probability and the economic fundamentals to get, by substituting one into the other, an estimable reduced form equation. Examples include Edwards (1986), Ozler (1993), Eichengreen and Mody (1999), Easton and Rockerbie (1999). Another solution has been to use proxies for the probabilities, like credit ratings (Kamin and von Kleist, 1999). A third approach has been to use multiple issues of the same borrower, assuming a common default probability (Cumby and Pastine, 2001). All these methods suffer from a common problem: they cannot identify more than one source of risk and test for a systematic extra effect of certain country characteristics.

Benczur (2007) suggests a rational expectations approach and proposes the errors in variables method (EVM) as a solution for these problems. Start from the representation of the structural form of the pricing equation:

$$s_{it} = \alpha + \beta R_t + \lambda E(d_{it}|Z_{it,}R_t) + \Theta X_{it} + \varepsilon_{1it}. \tag{1}$$

where s_{it} is the spread paid by country i on loans obtained at time t, R_t is the benchmark interest rate, d_{it} measures the loss from default, X_{it} is a vector of various extra factors, and Z_{it} contains information available at the time of pricing. Notice that Z_{it} contains all of the X_{it} variables as well. The error term ε_{1it} is orthogonal to any time t information (Z_{it} and R_t).

The linearity of (1) can be derived from risk-neutrality, profit maximization, and assuming partial default (with probability p) on the principal (denoted by x) but not on the interest

(r):

$$(1-p)(1+r)+p(x+r)=1+R,$$

which implies $s=r-R=p(1-x)$.

According to the EVM method, one replaces the expectation term in (1) with its realization:

$$d_{it}=E(d_{it}|Z_{it,}R_t)+\varepsilon_{2it}. \tag{2}$$

Given rational expectations, $E(\varepsilon_{2it}|Z_{it,}R_t)=0$, and thus equation (1) becomes:

$$s_{it}=\alpha+\beta R_t+\lambda d_{it}+\Theta X_{it}-\lambda\varepsilon_{2it}+\varepsilon_{1it}. \tag{3}$$

Now, d_{it} is not orthogonal to the compound error term, since it is not orthogonal to the prediction error ε_{2it} (see equation (2)) and ε_{1it} (possible simultaneity problem[3]). But according to the EVM approach, one can use the information set $\{Z_{it},\ R_t\}$ as valid instruments, since this set is correlated with the default event (from the prediction equation (2)) and uncorrelated with the error term (from the rational expectations assumption and the pricing equation (1)).

Actually, as Wickens (1982) argues, this method provides consistent though not fully efficient estimates even when the information set is incomplete or the functional form of the prediction equation is unknown. This is a major advantage, given the potential sensitivity of empirical results to functional form assumptions, selectivity bias and omitted variables. The key element is whether the fundamentals are sufficiently correlated with the default variable. If they are, then they can be used as valid instruments in the pricing equation, without having to specify the default prediction equation.

By using a set of instruments larger than the number of risk factors (which is one in our benchmark equation (1)), an overidentification situation arises. A rejection of this overidentification test could imply different conclusions. It might be that the rational expectations and risk neutrality assumptions are rejected. Or maybe the risk choice was the right one, but its indicator was imperfect. Our approach is that the instruments are proxying for a missing term in the structural equation, which is due to the direct effect of some factor on the spread, above the one coming from the influence on the predicted probability. In particular, if credit history has an extra punishment effect on top of influencing future repayment probabilities, it should show up as an extra determinant. Other factors like the

[3]This means the following. A pricing error ε_{1it} can lead to a higher future default loss d_{it}, the conditional expectation of which is a right hand side variable in the pricing equation. As the pricing error is assumed to be orthogonal to Z_{it} and R_t, it means that those variables remain valid instruments for this source of nonorthogonality as well.

level of reserves may also play a role.

To find such extra factors, there are two important points to check: first, whether a certain variable X_{it} has to be included as extra RHS variable for the overidentification test not to reject. In practice, this means looking at the increase in the p-value of the overidentification test brought about by the inclusion of certain extra terms. Second, whether the estimated Θ is significant. The two points ought to be connected, but it is possible that only one of the indications is present in small samples.

It is important to stress that a variable included in X_{it} affects the spread through two channels: through an impact on predicted future default $E(d_{it}|Z_{it,}R_t)$, and a direct effect through Θ. The total effect is captured by the reduced form equation

$$s_{it} = \alpha^{'} + \beta^{'} R_t + \Gamma^{'} Y_{it} + \Theta^{'} X_{it} + \varepsilon_{it}, \tag{4}$$

where $Y = Z \backslash X$. Denoting the linear conditional expectation of d_{it} by

$$d_{it} = \alpha^{''} + \beta^{''} R_t + \Gamma^{''} Y_{it} + \Theta^{''} X_{it} + \varepsilon_{2it},$$

the structural form (1) imposes the following restrictions on the reduced form (4):

$$\alpha^{'} = \alpha + \lambda\alpha^{''};\ \beta^{'} = \beta + \lambda\beta^{''};\ \Gamma^{'} = \lambda\Gamma^{''};\ \Theta^{'} = \Theta + \lambda\Theta^{''}. \tag{5}$$

This immediately shows the decomposition of the total effect $\Theta^{'}$ into the direct effect Θ and the indirect effect $\lambda\Theta^{''}$. Moreover, it also illuminates the way we identify this decomposition: $\Gamma^{''}$ and $\Theta^{''}$ are obtained from the prediction equation for d_{it}, the risk parameter λ is identified through the restriction $\Gamma^{'} = \lambda\Gamma^{''}$, while Θ is obtained as $\Theta = \Theta^{'} - \lambda\Theta^{''}$. This is exactly what an instrumental variables (IV) estimation of equation (1) does in one step. Finally, the overidentification test checks whether the matrix equation $\Gamma^{'} = \lambda\Gamma^{''}$ holds.

4 Data, variables and estimation issues

The choice of the time period was mostly driven by Ozler's (1993) observation that a period of market expansion is needed to distinguish the impact of an individual borrower's repayment history from the impact of a widespread panic. Thus, we use the period 1973-1981, which witnessed particularly intense syndicated bank lending to sovereign borrowers. In fact, bank loans were the dominant source of sovereign capital flows in the 70s, which was no longer true after the Debt Crisis. The initial dataset contains the spread (over the 1-year LIBOR) on

757 commercial bank loan contracts denominated in dollars, to 46 developing countries and were obtained from various issues of the World Bank's "Borrowing in International Capital Markets".[4]

As we have no access to contract-level characteristics of loans or their future repayment patterns, we average over all contracts of the same country at a given time period. Since the economic fundamentals are mostly available at the annually frequency, we construct yearly measures for the spread. Just like Easton and Rockerbie (1999), we use a weighted average of the original spreads. As an alternative, we also discuss results which use an average spread weighted by loan quantities only (like Edwards, 1986 and Ozler, 1993). Using maturity in the weighting allows for taking into account that the spread of a longer maturity debt influences average credit terms to a larger amount than the spread of a shorter maturity loan.[5] This transformation means that we are left with 252 yearly observations.

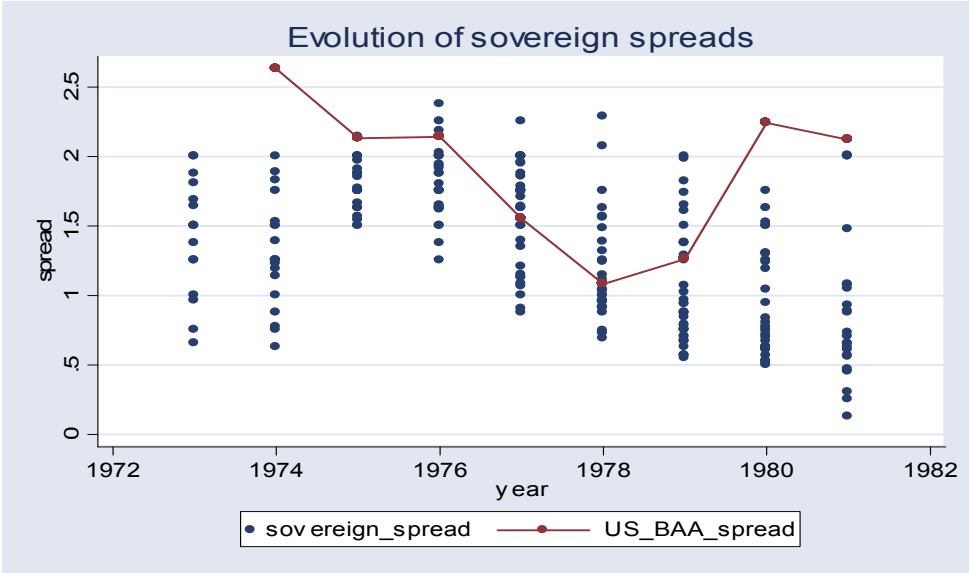
Data availability (arrears, country fundamentals) and the need of first differencing further reduce the working sample, to 157 observations from 37 countries. Finally, the smallest and largest values of our constructed variables are quite unrealistic, thus we excluded the top and bottom 1% for the recent default and 2% for the future default variable. The final sample for which we report the results is 149 observations and 36 countries.

Figure 1 illustrates the evolution of the sovereign spreads in the sample, together with the average BAA-rated US corporate bond spread (taken from the Federal Reserves' website). The spread is measured in percentage points. This comparison shows that sovereigns pay similar spreads to BAA-rated US companies. Another interesting aspect is that the variation in spreads is very large in 1981, suggesting that commercial banks were distinguishing between the borrowers, even before the "unexpected" debt crisis of 1982.

[4]Data on fees and commissions are not reliably available. It is noted, however, that these costs are low relative to the spreads (see Edwards, 1986).

[5]The World Bank's Global Development Finance database also reports average interest rates weighted by quantities. From a theoretical point of view, however, it is not obvious what the correct weighting should be. Suppose that a loan contract specifies an interest payment stream of a constant spread over the benchmark yield $r_t + s$ for a period T, at the end of which the full principal q is repaid (notice that actual contracts might specify a different schedule for principal repayments). It is straightforward to see that the present discount value (PDV) of such a loan contract is $qs[1 - \exp(-\int_0^T r_\tau d\tau)] = qs(1 - R_T^{-1})$. Here r_τ is the instantaneous required interest rate, while R_T is the required yield from time zero to T. Having two loan contracts (q_1, s_1, T_1) and (q_2, s_2, T_2), the equivalent contract $(q_1 + q_2, s', T')$ then must satisfy $q_1 s_1(1 - R_{T_1}^{-1}) + q_2 s_2(1 - R_{T_2}^{-1}) = (q_1 + q_2)s'(1 - R_{T'}^{-1})$. This defines s' for a given T'. In order for this to be a weighted average of s_1 and s_2, T' must be such that $R_{T'}^{-1}$ is the quantity weighted average of $R_{T_1}^{-1}$ and $R_{T_2}^{-1}$. Then s' is a weighted average of s_1 and s_2, with relative weights $q_1(1 - R_{T_1}^{-1})$ and $q_2(1 - R_{T_2}^{-1})$. For $T_1 \neq T_2$, this in general requires a detailed knowledge of the entire benchmark yield curve. If the yield curve is flat ($r_t = r$), R_{T_1} and R_{T_2} are not too large (in the sense that $R_{T_1}^{-1}$ and $R_{T_2}^{-1}$ are well approximated by $1 - R_{T_1}$ and $1 - R_{T_2}$), then the spread on the equivalent contract is precisely the quantity and maturity weighted average of s_1 and s_2.

Figure 1: The evolution of sovereign spreads for 1973-1981 (compared to the average BAA-rated US corporate bond spread; in percentage points)



4.1 Repayment history indicators

There is no clear indication from theory regarding the choice of the repayment history variable. Moreover, in the context of repeated games with imperfect monitoring, punishment is invoked by some imperfect indicator of cheating, and not an outright default episode. In general, the repayment history variable should still be related to the overall loss creditors incurred due to repayment problems. Recognizing that any indicator is merely a proxy, our choice is guided by data availability and explanatory power. In particular, binary indicators of repayment problems are available both historically and recently; arrears data is reported by the World Bank from 1971; while the size of debt forgiveness and reschedulings are reported by the World Bank only from 1989. For this reason, we use binary indicators for capturing "distant" repayment history and arrears data for "recent" history.

There are reasons to believe that recent and distant history have a different effect. Indeed, Ozler (1993) finds that repayment difficulties happening before the 1930s did not significantly matter for spreads in the 1970s, while those happening afterwards did. In the models of Yue (2006), Kovrijnykh and Szentes (2007), and Benjamin and Wright (2009), it is also recent default (arrears) that matters; in fact, once a country eliminates its arrears, it gets a clear credit history. From the point of view of relational contracts, we expect that it is mostly recent history that matters for potential punishments.

For distant history we use an indicator of the presence of default or rescheduling of bank loan debt to official creditors in the period 1940-1970. This dummy variable was constructed based on Ozler (1993), which includes data for 1956-1968, and Lindert and Morton (1989), which refers to the period 1940-1970. Table 1 presents its summary statistics, the number of observations and countries that had repayment problems. It is important to note that the indicator has significant variation to be able to identify the effects. The mean for this variable is high, and shows that around 30% of the countries in the sample had some repayment problems during 1940-1970.

Table 1: The "distant" default variable[a]

Variable: distant default dummy	Total observations /countries	Obs. with 1	Mean	Std. Dev.
Whole sample	252/46	70	0.335	0.473
Restricted sample	149/36	52	0.349	0.478

a Constructed as a dummy variable for repayment problems on loans for 1940-1970. The dummy takes the value 1 for a repayment problem.

While this indicator is very similar to those used in Ozler (1993), the indicators reflecting recent history are our own. Their construction allows including a continuous variable instead of a dummy to reflect past repayment behavior, enabling to control for country fixed effects and still include a default history measure. We construct this indicator from data on private arrears (both interest and principal) on long-term debt outstanding, available since 1971 from the Global Development Finance CD-ROM. As Cline (1984) notes, debt reschedulings are usually preceded by the accumulation of arrears, thus their presence and size can be a good indicator of potential creditor losses.

To reach this indicator, we add the change in the stock of arrears starting from the first lag and going back until 1971. The addition of these past values is weighted by linearly decreasing coefficients. This takes into account a "memory" that is fading in time, such that more distant events have a relatively smaller importance for the present than the more recent ones.[6] The Appendix contains a formal description of how we computed these indicators.

Finally, we further separate the effects of very recent history. One would expect that the indirect effect of a very close repayment problem (last 1-3 years) is more important than that of a more distant repayment problem, since it is very likely to reflect problems that generate additional repayment problems in the future. This should lead to a measure that is

[6] We also discuss results from additional ways of cumulating arrears: using a constant and a quadratic scheme.

significant in the reduced form but has no direct effect in the structural form. There might also be a "grace" period applied when the banks decide whether to charge higher spreads for defaulting countries, in which case very recent history might not matter. Thus, we separate the ten years measure into the last three years and the preceding seven. The first choice, however, is not significant even in the reduced form, leading us to believe that there was indeed a grace period attached to previous defaults. The variable that adds the flows of arrears from time t-4 to t-10, on the other hand, is significant. Thus it is the one we report in our results section.

We would stress again that there is no clear indication from theory regarding the choice of the repayment history variable. In particular, one could also employ a similar measure based on arrears normalized by the amount outstanding. Such an indicator, however, does not have plausible reduced form estimates: as discussed in section 6.3, its parameter is usually negative and/or insignificant. Being guided by data availability and explanatory power, we take this as an indication that the absolute measure is more successful in capturing the effect of repayment history.

4.2 Future default variables

Unlike past default indicators, the future default variable should closely reflect the realization of proportional losses on a loan. As demonstrated by Sturzenegger and Zettelmeyer (2007), and Benjamin and Wright (2009), precise measures of realized repayments are very hard to compute for sovereign debt. Easton and Rockerbie (1999) argue that arrears are more indicative for repayment problems than default or rescheduling indicators. Based on these, we construct our future default measure by using again GDF data on arrears: we add private arrears for a period similar to the average maturity. As opposed to our measure for recent default, there is no ambiguity that this sum needs to be normalized by the total loan amount corresponding to that observation, in order to reflect relative losses. The detailed procedure is explained in the Appendix.

Table 2 provides some brief descriptive statistics of our benchmark choice for future and recent default. For recent default around 65% of the observations are equal to 0 for both the full and the reduced sample; this number is around 20% for the future indicator. The difference is due to more frequent arrears after 1981, but it shows that there were still countries that were not accumulating arrears in this period.

Table 2: The "recent" and "future" default variables[a]

Variable Variable	Total obs. /countries	Obs. with 0	Mean	Std. Dev.	10%	Median	90%
Recent default[b]	211/40	155	3.212	22.996	-0.286	0	0.2
Restricted sample	149/36	102	1.187	5.829	-0.0571	0	0.229
Future default[c]	207/39	43	0.376	2.115	0.003	0.0007	0.461
Restricted sample	149/36	28	0.207	0.741	-0.0096	0.0009	0.481

a Constructed as continuous variables based on arrears data. A zero means no repayment problem.
b The indicator adds private arrears for the time t-4 to t-10 and uses linearly decreasing weights. Information refers to the whole sample.
c The indicator adds private arrears for 8 years in the future and divides them by the loan amount. Information refers to the whole sample.

4.3 Economic fundamentals

The sources for these variables are Global Development Finance, International Financial Statistics and World Development Indicators. Besides data availability, we are following most of the literature in considering the following variables as candidates: *debt to GDP, reserves to imports, debt service to exports, current account per GDP, exports to GDP, savings to GDP, growth of per capita GDP, growth of gross investment, GDP per capita, inflation, credit to private sector per GDP.*

We construct two additional variables that are related to the international financial environment of a country. *Repeated borrowings* are designed to capture the importance of relationship banking. It is constructed by cumulating the number of months in which the borrower received loans. *Proportion of countries with arrears in the region* captures a regional contagion effect from one country going into arrears. It is obtained by dividing the number of countries with arrears from the same region by the total number of countries in that region. Finally, for the benchmark interest rate we use the LIBOR USD 1-year rate.

4.4 Estimation issues

Both the reduced and the structural form specification are subject to two major econometric problems: the need to control for country fixed effects, and the validity of the strict exogeneity assumption.[7] The Appendix contains an in-depth discussion of how we handle these issues,

[7]Strict exogeneity means that the idiosyncratic error terms, conditional on the individual effect, are uncorrelated with past, present and future values of the regressors.

here we just briefly outline our strategy.

In the reduced form, we use a pooled OLS, a random effects and a fixed effects indicator. As none of them are appropriate when strict exogeneity fails, our preferred specification is first differencing and instrumenting the 'proportion of countries with arrears' variable with its first lag. In the structural form, we use the first difference estimator with appropriately instrumenting the future default variable: it eliminates the individual effects, and the right choice of instruments resolves the endogeneity problem caused by the prediction error. The appropriate instruments include the first and/or second lags of the regular instruments (time t information). Using as instruments the levels of the variables (as opposed to the second difference) leads to more precise estimates, but at the cost of making the direct comparison with the reduced-form results of Table 3 more difficult.

The whole structural estimation framework is based on the validity of the instruments: they should be correlated with the instrumented variable and also uncorrelated with the error terms. For this reason, we report several measures that summarize the first-stage regression, and an overidentification test that is appropriate in a setting with heteroskedasticity and autocorrelation (the Hansen's J statistic), and it is consistent even with intra-cluster correlation.[8]

5 Results

5.1 The reduced form

The results of the reduced-form estimation are presented in Table 3 and refer to four specifications. The left hand side variable is a quantity and maturity weighted annual spread. The recent default variable used is the year $t-4 \ldots t-10$ private arrears flow, weighted by linearly decreasing weights. As Column (1) shows, most of the explanatory variables are significant and have the expected sign in the pooled OLS specification. In particular, the coefficient of both distant and recent default is positive and significant (columns 1, 3 and 1-4, respectively). When running panel data specific regressions like fixed effects (Column 2) and random effects (Column 3), the results do not change substantially except for two variables, savings to GDP and experience on the market.

Column (4) is a specification that uses only the sequential exogeneity assumption: the variables are first-differenced and the proportion of countries with arrears is instrumented with its first lag.[9] The estimates' precision is smaller, due to the fact that there is less time-

[8]This feature is important because Baum et al.(2003) cite evidence that show that the presence of intra-cluster correlation can readily cause a standard overidentification statistic to over-reject the null.

[9]The difference with the not-instrumented FD version is mostly with respect to the significance of this

variation and that by first differencing there is one observation not used for each country. Still, the effect of recent default is positive and (marginally) significant, even after controlling for country effects.

In conclusion, our findings suggest that fixed effects do play a role, that strict exogeneity is an unreasonable assumption for most of the variables (invalidating the classical panel data methods) but after controlling for these issues there are economic fundamentals, including recent default, that significantly influence the spread. There is also some indication that the significance of distant default is not just a consequence of omitted country effects.[10]

5.2 The structural form

The results of our benchmark specifications are presented in Table 4. As motivated in Section 5, we choose as our main structural form estimation the one in first differences, since it provides the framework that allows us to make correct inference on the overidentification test and the RHS variables. We also briefly comment on some results from the level specification, as they can shed light on the channel decomposition of distant default.

Overall, there are three important findings we discuss here: the influence of the future default indicator, the coefficient of the benchmark yield, and most importantly, the effect of the recent default indicator. Starting with the first, the future default's point estimate is strikingly robust across all methods (being around 0.35), and in almost all of them significant at the 10% level.

Although the mean of this indicator is just 0.2, this is not very indicative of its influence, because the variance is large and for many countries the indicator's value is around 0.5 and even 1. If we consider an increase in the indicator from its median to the 90th percentile, then this would raise the spread by approximately 0.17. Consequently, the coefficient can be considered as sizable, as the sample mean of the spread is 1.33. This is an important finding, because it suggests that expected default risk was priced in the lending decision and that the debt crisis of 1982 was, to this extent, "anticipated". Nevertheless, it remains true that for many countries this prediction was correct mainly in sign, and less in the size, as the quantitative effect of the future default risk is small.

variable.

[10]Though fixed effects are not compatible with our distant default variable, the comparison of the distribution of country effects across groups with different default history might convey some information. Indeed, the results from the reduced-form estimations suggest that defaulters (categorized through the distant default dummy) are being charged significantly higher spreads. The difference between the means of the two groups is 0.3, which is around one and half times the estimate on the default dummy (see Table 3). Moreover, the standard deviation of the country fixed effects is 0.3; hence it could be argued that the estimated default dummy explains around 70% of the typical variation of the country fixed-effects.

Table 3: Reduced-form estimation: the determinants of the spread[a]

	Pooled OLS	Fixed Effects	Random Effects	First Difference [b]
	(1)	(2)	(3)	(4)
Benchmark yield	−0.097	−0.096	−0.094	−0.14
	(−10.78)**	(−9.42)**	(−11.36)**	(−2.92)**
Distant default	0.22	–	0.19	–
(1940-1970)	(2.54)**		(1.86)**	
Recent default	0.0097	0.017	0.014	0.0099
(last 10 years)	(1.67)*	(1.80)**	(1.68)**	(1.48)*
Reserves to imports	−0.84	−0.66	−0.73	−0.79
	(−4.96)**	(−3.41)**	(−5.04)**	(−2.93)**
Exports to GDP	0.80	1.11	0.81	1.37
	(2.47)**	(2.08)**	(2.61)**	(2.02)**
Savings to GDP	−0.70	−0.69	−0.83	−0.74
	(−1.26)	(−1.34)*	(−2.14)**	(−1.07)
Repeated borrowings	−0.0017	−0.0048	−0.0037	−0.57
	(−0.61)	(−1.66)**	(−1.34)*	(−1.72)**
Countries with arrears	0.50	0.88	0.55	4.98
(% of region total)	(2.06)**	(1.81)**	(2.41)**	(1.42)*
Constant	2.26	2.06	2.23	–
	(11.66)**	(9.10)**	(14.57)**	
No. of obs.	178	178	178	149
p-value[c]	0.0000	0.0000	0.0000	0.0000

a The t statistics are in parentheses; the standard errors are corrected for clustering at country level. *, ** denote 0.2 and 0.1 significance levels, respectively.

b The proportion of countries of arrears (first differenced) is instrumented with its first lag.

c The p- value of joint significance of the regressors.

Table 4: Structural-form estimation: the determinants of the spread[a]

	Estimation method [b]		
	(1)	(2)	(3)
Future default	0.361	0.378	0.352
	(2.26)**	(2.29)**	(2.28)**
Benchmark yield	−0.100	−0.099	−0.095
	(−10.52)**	(−9.98)**	(−9.92)**
Recent default (last 10 years)		0.0291	0.0156
		(5.26)**	(3.10)**
Reserves to imports			−0.703
			(−2.95)**
First stage relevance:[c]			
R^2 for future default	0.089	0.099	0.104
F statistics[d]	16.83	9.18	6.76
Anderson Canon. Corr. LR p-val[e]	0.03	0.01	0.01
Structural form:			
Overidentification test p-value[f]	0.198	0.327	0.463
Number of observations	149	149	149

a The t statistics are in parentheses; the standard errors are corrected for clustering at country level. . *, ** denote 0.2 and 0.1 significance levels, respectively.

b The dependent and explanatory variables are first differenced, while the instruments are in general first lags of their levels.

1. The future default variable is instrumented by the first lag of the following variables: benchmark yield, reserves to imports, savings/GDP, arrears in the region, experience on the market, exports/GDP, and recent default.
2. As in 1, but including the recent default (first differenced, not instrumented) as extra exogenous RHS variable.
3. As in 2, but including reserves/imports (first differenced, not instrumented) as extra exogenous RHS variable.

c The reduced form regression of the instrumented indicator(s) on the full set of instruments.

d From the F test of the joint significance of the excluded instruments in the first-stage regression.

e The Anderson Canonical Correlations Likelihood Ratio test of the null hypothesis that the equation is underidentified.

f The Hansen J-statistic .

A second general feature is the stability and significance of the benchmark rate, which was also present in the reduced-form results. Actually, the most robust and significant result of both the reduced and structural form is the negative coefficient on the benchmark yield. Given that the spread is defined as the difference between the loan rate and the LIBOR rate, this result is equivalent to the finding that the loan rate responds less than one-in-one to the world interest rate (the reaction coefficient is around 0.9). This conclusion is found also in Eichengreen and Mody (1999) and Uribe and Yue (2006).

The central results concern the channels of influence of the economic fundamentals, particularly that of the recent default indicator. The starting point is to use the benchmark yield and the future default indicator as the only explanatory variables, which is done in Column (1). The first-stage summary statistics indicate that the instruments are relevant, with a partial R^2 of 0.089 and a value for the F test of the joint significance of the excluded instruments of 16.83.[11] Also the Anderson canonical correlations likelihood-ratio test of whether the excluded instruments are relevant (the null being that the equation is underidentified) rejects the null with a p-value smaller than 0.04.

However, the overidentification test does not offer strong evidence of a correct specification, with the p-value being 0.198. We interpret this as indicating that the influence of a subset of our instruments on the spread does not go only through predicted default probability. The next step is to look for variable(s) that could improve upon overidentification and/or be significant as extra RHS variables. Although these two criteria should give the same conclusion, the small sample implications might affect their power in detecting the right answer. The natural candidate is the recent default variable. Its inclusion gives the results displayed in Column (2): the p-value of the overidentification test increases to 0.327, and the variable is very significant and has the right sign.

Based on the p-value, this might be accepted as the final specification. To be on the safe side, however, we still explore whether the inclusion of some additional variables would further improve upon overidentification. After a thorough search, we find that the reserves to imports ratio is such an additional variable: it has a positive, significant and robust coefficient and it also improves the overidentification result (the p-value becoming 0.463, as in Column (3)). It appears then that both these extra RHS variables improve on the overidentification test. Thus, after running several robustness checks we interpret Column (3) as our benchmark specification for the structural form.

The recent default measure is not very sizeable. Still, our analysis shows that this measure is statistically significant and in some cases economically too, as there are countries for which

[11] This value is higher than the one, equal to 10, considered in Staiger and Stock (1997) as a rule of thumb critical value for considering the instruments weak.

this extra punishment effect accounts for 32-56 basis points.[12] Moreover, the punishment lasts for more than one period: based on our benchmark choice of the recent default indicator, going into arrears raises the spreads four to ten years after the episode itself, though the impact is diminishing in time. As Tedeschi (1994) suggests, even a punishment that is relatively small in each period can deter deviations (in his case: cheating in the cartel), as long as it can last for a potentially very long time. As an extreme case, if the punishment threat is sufficient to deter default completely, then one would not observe any punishment at all. Consequently, even a small price punishment is indicative for the presence of relational contracts.

As a specific example, consider the following hypothetical dilemma of a country. It has a preexisting stock of arrears of 61.9 million dollars, and it can expect an annual flow of new bank loans of 3861 million dollars (the figures correspond to Brazil in 1977). According to our results, eliminating all of its arrears would lead to smaller spreads on future bank loans: starting four years later, the spread would be lower by 68 basis points, then 58 basis points etc., and the impact would finally disappear by the eleventh year. Let us compare three alternatives: (1) clearing all arrears now, (2) clearing them next year, (3) rolling them over indefinitely. Using the conventional 10% discounting, the difference in the present value of repayment costs between scenario 1 and 2 is 5.4 million dollars, while the difference between scenario 1 and 3 is 59.7 million dollars. This latter number is already quite comparable to the amount of arrears in consideration.

Finally, we briefly touch upon the level specification. The only reason why this is relevant is that it can separate the effect of distant default by channels. Unfortunately, estimation in levels is plagued with several problems which render it unreliable. The presence of fixed effects calls for a FE or RE specification, where the strict exogeneity assumption fails, and it suffers from a very significant rejection of the overidentification test. The instrumented future default risk appears very insignificant and the point estimate is much lower than the one resulting from the correctly specified first differenced estimation. We attribute these to the failure of the strict exogeneity assumption.

If instead we ignore fixed effects and run the pooled level IV the results are broadly close to our benchmark first-difference specification: the instruments have a similar first-stage explanatory power, there are extra structural effects coming from recent default and reserves to imports ratio, with the latter two variables also improving on the overidentification test up to 0.2. When the past default dummy is included as an extra RHS variable, it is not

[12] For the reserves to imports ratio variable, the quantitative effect is more pervasive. An increase in the measure from its median to its 90th percentile would increase the spread by 24 basis points (around 20% of the average spread).

significant and decreases the p-value on the overidentification test. Thus, this specification would suggest that past distant default affects the spread through future default risk, and not through a punishment channel. Given the problems outlined earlier, we would not trust enough this specification to make a strong conclusion.

5.3 Robustness

While checking the robustness of our results, we found that many initial specifications were sensitive to a very small number of extreme outliers. The results presented here are, on the contrary, robust to most changes in the set of observations or instruments. The null of homoskedasticity in the regression of Column (3) is not rejected by the Pagan-Hall general test statistic (p-value of 0.49), which suggests that, given its inferior small sample properties, an asymptotically efficient GMM is not needed.[13] There is also no evidence for serial correlation of the errors in this first difference specification.

We now discuss several alternatives in the construction of three key variables: the annual average spread, the recent and the future default variable. We first check whether these alternative default indicators are good proxies (in the sense of having a significantly positive reduced form effect),[14] and then for those which pass this "reduced form test" we check whether our benchmark results remain valid. In all cases, we followed the same outlier filtering applied to the particular default indicator at hand (excluding the top and bottom 1% for the recent default and 2% for the future default indicator).[15]

Starting with the first, we reran our estimations (the ones reported in tables 3 and 4) with a quantity-weighted average spread (instead of weighting by quantity and maturity). Reassuringly, the point estimates, their significance level, and the evolution of the overidentification test's rejection level have remained almost identical.

For the recent default measure, we varied (1) the discounting scheme (besides linear, we worked with a quadratic and a constant scheme), (2) the time window (besides $t-4 \ldots t-10$, we also used $t-1 \ldots t-10$ and $t-1 \ldots t-3$), and (3) whether it is an absolute or relative measure (in dollars or in proportion to some measure of the stock of debt). Quadratic discounting had some influence on the first stage fit and the overidentification test of the structural form. In particular, the partial F has increased and remained above 19 for all three cases, while the inclusion of recent default does not increase the p value of the overidentification test. As the point estimate of recent default remains highly significant and positive, this specification also confirms our benchmark estimates. Turning to the version

[13] Even so, results from a GMM regression were found to be very close.

[14] In all cases, the reduced form effect of distant default remained similar in size and significance.

[15] Detailed results are available from the authors upon request.

without discounting, its reduced form parameter becomes insignificant, with the exception of the FE method; implying that it is not a successful proxy for capturing the total effect of recent default.

Next we vary the time window for which we add up past arrears, combined with linear or no discounting. Again, these measures do not pass the reduced form test: with the exception of pooled OLS, the point estimates are insignificant and/or negative. Finally, we consider versions when we add up past arrears normalized by the amount of new loans of the corresponding year. Unfortunately, these proxies also fail the reduced form test, yielding highly insignificant and often negative point estimates.

Regarding our future default indicator, we consider two alternatives: in one case we split future arrears evenly among past contracts (instead of splitting in proportion of loan size, see equation (6) in the Appendix), and in the other we exclude a three year grace period from the summation in equation (6). The first variant has a much lower first stage fit, and the point estimate of future default is also much less significant. This suggests that this indicator is less able to represent default risk in our structural equation. A third difference is that the inclusion of recent default alone does not improve on the overidentification p value, though its size and significance is similar to the benchmark. Finally, allowing for a three year grace period leads to very similar estimates and conclusions, with a more modest first stage fit.

5.4 The channel decomposition

An interesting exercise is to compare the point estimates from the reduced and structural form estimation for the extra risks (recall equation (5)). For the reserves variable, the estimated coefficient in the structural form is about 90% of that from the reduced form. We interpret this as evidence that this factor has some effect on the spread going through the expected default risk, but most of the influence is through a direct channel. For the recent default, the same comparison leads to the conclusion that its effect is almost entirely a direct one, pointing to a price punishment above increased future default risk. In other words, there is evidence for positive lender surpluses after a default episode, indicating the presence of relational contracts in sovereign bank lending.

Regarding the structural effect of reserves, we offer three interpretations. The first involves a systematic pricing error: either banks systematically overreacted to movements in reserves, or the countries themselves chose, ex post, to invalidate the expectations and defaulted less than anticipated. The second explanation views the reserves to imports ratio as a proxy for a currency or liquidity crisis risk. Such an event might imply a contagion from

the local currency market to foreign currency markets, or financial market liquidity may decline, making banks also reluctant to lend at unchanged terms. The third explanation views an increase in reserves as a signal of potential cheating. If a country is contemplating a default, it anticipates limited future access to borrowing, thus it starts piling up extra reserves. Understanding this, creditors may invoke a price punishment.

6 Concluding remarks

We extended the existing empirical literature on the role of credit history in sovereign bank lending along two dimensions. One is that we used a continuous measure of past default, enabling us to control for country fixed effects. Our other, more important contribution is the empirical strategy that allows for the distinction of multiple channels of influence. This strategy is a structural asset pricing rational-expectations estimation in which the spread may be influenced by multiple risks and factors. Using the errors in the variables method, we replace the expectation term with its realization and instrument the latter with information available at the time of pricing. Using the overidentification test, we then investigate whether there is any instrument that should be added as an extra RHS variable. We interpret any such variable as influencing the spread not only through expected default risk, but having an extra effect on it.

The reduced-form estimation provides evidence that, after controlling for fixed effects, borrower and regional characteristics, both recent and distant repayment history are significant. This makes the result similar to that obtained by Ozler (1993) and implies that although country effects do matter, credit history does play a role in determining sovereign spreads. The structural-form regression provides strong evidence of an extra effect of credit history (a punishment effect) in *prices*, above the one going through predicted default loss. The finding that credit history matters beyond predicting future default points to the presence of relational contracts in sovereign bank lending, where repayment incentives are incorporated into future borrowing terms. The major structural specification includes the benchmark LIBOR interest rate, expected default risk, the recent default indicator and the reserves to imports ratio. All these variables are significant and robust to different specifications.

In terms of the default costs, we do believe that in reality there is a complex mix of trading and political sanctions, spillovers to other transactions and relationships, signaling and reputation considerations (self enforcing contracts). Our main result is that there is evidence of this last effect: an extra surcharge in loan prices. This is consistent with the presence of relational contracts in sovereign bank lending.

References

Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a theory of discounted repeated games with imperfect monitoring," *Econometrica*, 1041–1063.

Arellano, C. (2008): "Default risk and income fluctuations in emerging economies," *American Economic Review*, 98, 690–712.

Baker, G., R. Gibbons, and K. Murphy (2002): "Relational Contracts and the Theory of the Firm*," *Quarterly Journal of Economics*, 117, 39–84.

Baum, C. F., M. E. Schaffer, and S. Stillman (2002): "Instrumental variables and GMM: Estimation and testing," Boston College Working Papers in Economics 545.

Benczur, P. (2007): "Identifying Sovereign Bond Risks," Central European University Working Paper 9.

Benjamin, D. and M. Wright (2009): "Recovery Before Redemption: A Theory of Delays in Sovereign Debt Renegotiations," Mimeo.

Chatterjee, S., D. Corbae, M. Nakajima, and J. Ríos-Rull (2007): "A Quantitative Theory of Unsecured Consumer Credit with Risk of Default," *Econometrica*, 75, 1525–1589.

Cline, W. (1984): *International Debt: Systemic Risk and Policy Response*, Washington D.C.: Institute of International Economics.

Cumby, R. E. and T. Pastine (2001): "Emerging market debt: measuring credit quality and examining relative pricing," *Journal of International Money and Finance*, 20, 591–609.

Easton, S. T. and D. W. Rockerbie (1999): "What's in a default? Lending to LDCs in the face of default risk," *Journal of Development Economics*, 58, 319–332.

Eaton, J. (1996): "Sovereign Debt, Reputation and Credit Terms," *International Journal of Finance & Economics*, 1, 25–35.

Eaton, J. and M. Gersovitz (1981): "Debt with Potential Repudiation: Theoretical and Empirical Analysis," *Review of Economic Studies*, 48, 289–309.

Edwards, S. (1986): "The pricing of bonds and bank loans in international markets : An empirical analysis of developing countries' foreign borrowing," *European Economic Review*, 30, 565–589.

Eichengreen, B. and A. Mody (1999): "Lending Booms, Reserves, and the Sustainability of Short-Term Debt: Inferences from the Pricing of Syndicated Bank Loans," NBER Working Papers 7113, National Bureau of Economic Research, Inc.

Green, E. and R. Porter (1984): "Noncooperative collusion under imperfect price information," *Econometrica*, 87–100.

Greif, A. (1993): "Contract enforceability and economic institutions in early trade: The Maghribi traders' coalition," *The American Economic Review*, 525–548.

Hopenhayn, H. and I. Werning (2008): "Equilibrium default," Manuscript, MIT.

Kamin, S. B. and K. von Kleist (1999): "The evolution and determinants of emerging market credit spreads in the 1990s," International Finance Discussion Papers 653, Board of Governors of the Federal Reserve System (U.S.).

Keane, M. P. and D. E. Runkle (1992): "On the Estimation of Panel-Data Models with Serial Correlation When Instruments Are Not Strictly Exogenous," *Journal of Business & Economic Statistics*, 10, 1–9.

Kletzer, K. and B. Wright (2000): "Sovereign Debt as Intertemporal Barter," *American Economic Review*, 621–639.

Kovrijnykh, N. and B. Szentes (2007): "Equilibrium Default Cycles," *Journal of Political Economy*, 115, 403–446.

Levenstein, M. (1997): "Price wars and the stability of collusion: a study of the pre-World War I bromine industry," *The Journal of Industrial Economics*, 117–137.

Lindert, P. H. and P. J. Morton (1989): "How Sovereign Debt Has Worked," in *Developing Country Debt and Economic Performance, Volume I: The International Financial System*, ed. by J. Sachs, University of Chicago Press.

Milgrom, P., D. North, and B. Weingast (1990): "The Role of Institutions in the Revival of Trade: the Law Merchant, Private Judges, and the Champagne Fairs," *Economics & Politics*, 2, 1–23.

Obstfeld, M. and K. Rogoff (1996): *Foundations of International Macroeconomics*, MIT Press.

Ozler, S. (1993): "Have Commercial Banks Ignored History?" *American Economic Review*, 83, 608–20.

Panizza, U., F. Sturzenegger, and J. Zettelmeyer (2009): "The Economics and Law of Sovereign Debt and Default," *Journal of Economic Literature, forthcoming.*

Porter, R. (1983): "A Study of cartel stability: The joint executive committee, 1880-1886," *The Bell Journal of Economics*, 301–314.

Reinhart, C. M., K. S. Rogoff, and M. A. Savastano (2003): "Debt Intolerance," *Brookings Papers on Economic Activity*, 34, 1–74.

Sandleris, G. (2008): "Sovereign defaults: Information, investment and credit," *Journal of International Economics*, 76, 267–275.

Staiger, D. and J. H. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65, 557–586.

Sturzenegger, F. and J. Zettelmeyer (2007): "Creditors' Losses Versus Debt Relief: Results from a Decade of Sovereign Debt Crises," *Journal of the European Economic Association*, 5, 343–351.

Tedeschi, P. (1994): "Cartels with homogeneous and differentiated products and imperfect monitoring," *International Economic Review*, 635–656.

Uribe, M. and V. Z. Yue (2006): "Country Spreads and Emerging Countries: Who Drives Whom?" *Journal of International Economics*, 69, 6–36.

Wickens, M. R. (1982): "The Efficient Estimation of Econometric Models with Rational Expectations," *Review of Economic Studies*, 49, 55–67.

Wooldridge, J. (2002): *Econometric Analysis of Cross Section and Panel Data*, MIT Press.

Wright, M. L. (2002): "Reputations and Sovereign Debt," Stanford University. Mimeo.

Yue, V. Z. (2006): "Sovereign Default and Debt Renegotiation," Mimeo.

Appendix

Appendix A: Construction of the recent default indicator

1. For each country we use the time period 1971-1981 and for each of these years, we compute the flows of new private arrears as the first difference of the stock of arrears (for 1971, it is the stock itself). Let this measure be denoted by $arrears_{jt}$, where j=1,...,36 for countries and t=1,...,11 (corresponding to the period 1971-1981).

2. For the same balanced panel coverage we compute the following indicator:

$$recentdef10_{jt} = \sum_{k=1}^{t} arrears_{j(t-k)} \left(1 - \frac{k-1}{10}\right)$$

for each country j =1,...,36 and each t=1,...,11. Thus it is an indicator for the discounted repayment problems of the last 10 years.

Appendix B: Construction of the future default indicator

Let us denote the yearly GDF data on arrears by $A_{i,s}$, with s ranging from 1970 to 1989. As those arrears may refer to contracts from multiple years, we need to allocate them among contracts. Our approach involves the following steps. First, we split $A_{i,s}$ among contracts from years s-1,s-2,...,s-8 in proportion to new disbursements in the corresponding year:

$$Y_{is,s-j} = A_{i,s} ND_{i,s-j}/(ND_{i,s-1} + ... + ND_{i,s-8}) \tag{6}$$

$j = 1...8$ (the average loan maturity in our sample). Here $Y_{is,s-j}$ is the time s arrears on a time $s-j$ contract and $ND_{i,s-j}$ is the size of the time $s-j$ loan for country i. Then we cumulate the arrear fragments $Y_{it+1,t}, Y_{it+2,t}, ..., Y_{it+8,t}$ into which a time t contract goes over its lifespan:

$$D_{it} = Y_{it+1,t} + Y_{it+2,t} + ... + Y_{it+8,t} \tag{7}$$

Thus, in order to recover the amount of arrears affecting an active loan contract in year t, we assume that all the time t arrears affect all the loans that have not matured yet, and the size of the contract specific arrear is proportional to the size of the contract. We motivate this by two arguments: one is that there is no information available on which contracts these arrears correspond to; and second, the assumption that these flows can be attributed to several preceding loans is consistent with the cross-default clauses that these contracts included. According to such clauses once a country enters into default or any repayment problem that constitutes a break on the contract with one lender, this will be treated as

default also by the other creditors. Alternatively, instead of using loan disbursements as weights, we discuss results with equal weights. We also consider the effect of the grace period, thus excluding from this summation contracts that are more recent than the average grace period of 3 years.

Appendix C: Estimation issues

The reduced form in a panel framework is:

$$s_{it} = \alpha + \beta R_t + \Gamma Z_{it} + c_i + \varepsilon_{it},$$

where ε_{it} is the idiosyncratic error term, Z_{it} are the economic fundamentals for country i known at time t and c_i is the unobservable individual effect.

The first major concern is that the usual pooled OLS estimates are incompatible with individual country effects. The usual procedure to correct for fixed effects is a fixed effects (FE) or a random effects (RE) estimator. A key assumption behind both methods is strict exogeneity, which requires that the idiosyncratic error terms, conditional on the individual effect, are uncorrelated with past, present and future values of the regressors. If this fails, then all the classic panel data methods and specification tests are inconsistent. Formally, the strict exogeneity assumption means: $E(\varepsilon_{it}|Z_{is}, c_i) = 0$, for all t and s. There are reasons to suspect that the assumption might fail, as any pricing error (ε_{it}) could affect the future values of certain indicators, like reserves, debt to GDP, participation on the market, proportion of countries in arrears.

Wooldridge (2002) suggests a test for this: use the FE estimator but also include future values of some variables that are likely to break the assumption. Their significance is evidence that the assumption is likely to fail, at least for those variables. When performing such a test, we do actually find that when we include leads of variables such as: proportion of countries with arrears, experience on the market, saving to GDP, debt to GDP, they are significant.[16]

The strict exogeneity assumption is even more problematic and crucial in the structural form (equation (1) with the country effects included) than in the reduced form. As Keane and Runkle (1992) strongly point out, in this type of models, there are never any strict exogenous variables or instruments. This formal result comes from the effect of the prediction error on the future values of the variables.

When there are concerns about strict exogeneity, the general approach is to use a

[16]The strict exogeneity assumption is likely to be satisfied for our recent default variable since it captures arrears up to time t-4, thus it is insulated from the present time t through time t+3 pricing error.

transformation to remove the country effects c_i, and then search for instrumental variables, assuming only sequential exogeneity (Wooldridge, 2002). According to this assumption, the idiosyncratic errors, conditional on c_i, should be uncorrelated with the contemporaneous and past values of the regressors (instruments), but not with future values.

In this respect, a first-difference (FD) estimator is attractive:

$$s_{it} - s_{i(t-1)} = \beta(R_t - R_{t-1}) + \Gamma(Z_{it} - Z_{i(t-1)}) + \varepsilon_{it} - \varepsilon_{i(t-1)}.$$

One can notice that if strict exogeneity fails, then there is a problem here as well, since $E(\varepsilon_{i(t-1)}|Z_{it}) \neq 0$. The sequential exogeneity assumption, however, implies that all the lags of Z (or their linear combination) can be used as potential instruments for $Z_{it} - Z_{i(t-1)}$ and then the estimation is consistent.

For the structural form, this would mean estimating:

$$s_{it} - s_{i(t-1)} = \beta(R_t - R_{t-1}) + \lambda(d_{it} - d_{i(t-1)}) + \Theta(X_{it} - X_{i(t-1)}) - \lambda(\varepsilon_{2it} - \varepsilon_{2i(t-1)}) + \varepsilon_{1it} - \varepsilon_{1i(t-1)}.$$

However, the first difference specification causes further complications. The rational expectation assumption guarantees that the prediction error ε_{2it} is orthogonal to time t information, but this is not true about $\varepsilon_{2i(t-1)}$. The remedy is to use $Z_{i(t-1)}$ or $Z_{i(t-2)}$ as instruments, as those variables are not correlated with any error at time t or $t-1$.

Including extra RHS variables one could also worry about the possible failure of strict exogeneity for these variables. However, since in the benchmark specification these extra variables will be reserves to imports and recent default, for which we do not have reduced form evidence about the lack of strict exogeneity, we do not instrument them. Moreover, not instrumenting the extra RHS variables allows for a clearer and more direct comparison with the reduced form regression. Thus, whenever we refer to instrumenting in the structural-form we mean instrumenting only the future-default variable.

Appendix D: Tables not for publication

Table 5: The reduced form test (not for publication)

		Specification									
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		benchmark	weighting	discounting		time window				relative	
Pooled OLS	point estimate	0.01	0.01	0.011	**0.006**	0.006	0.005	0.006	0.006	**-12.35**	**-4.34**
	(p value)	0.103	0.108	0.039	**0.304**	0.04	0.093	0.133	0.112	**0.207**	**0.844**
FE	point estimate	0.017	0.016	0.014	0.016	**0.027**	**0.017**	**-0.004**	**-0.000**	**-6.98**	**47.19**
	(p value)	0.074	0.087	0.137	0.076	**0.761**	**0.145**	**0.658**	**0.979**	**0.723**	**0.317**
RE	point estimate	0.014	0.013	0.014	0.011	**0.006**	**0.008**	**0.003**	**0.004**	**−8.11**	**6.64**
	(p value)	0.092	0.103	0.102	0.142	**0.227**	**0.176**	**0.618**	**0.482**	**0.626**	**0.815**
FD	point estimate	0.01	0.01	0.008	**0.014**	**-0.014**	**-0.005**	**-0.011**	**-0.007**	**-63.8**	**15.26**
	(p value)	0.148	0.164	0.191	**0.375**	**0.148**	**0.743**	**0.173**	**0.637**	**0.153**	**0.819**
Reduced form test		*pass*	*pass*	*pass*	*fail*	*fail*	*fail*	*fail*	*fail*	*fail*	*fail*
Sample size	level	178	178	177	176	177	175	173	171	174	177
	FD	149	149	149	147	147	145	146	143	146	149

All columns report the reduced form coefficient and the p value of recent default. Column 1 is our benchmark, where we use (1) a quantity and maturity weighted annual spread, (2) linear discounting, (3) a time $t-4 \ldots t-10$ time window and (4) absolute arrears in constructing our recent default indicator. Column 2 uses only quantities as weights in constructing the annual average spread. Column 3 uses a quadratic, while column 4 uses no discounting. Column 5 uses a $t-1 \ldots t-10$ time window with linear discounting, while column 6 uses no discounting. Column 7 uses a time $t-1 \ldots t-3$ time window with linear discounting, while column 8 uses no discounting. Column 9 and 10 add up past arrears normalized by the amount of new loans of the corresponding year, for a $t-1 \ldots t-10$ and a $t-4 \ldots t-10$ time window (and linear discounting).
Numbers in bold indicate the reason why we classified the version as failing the reduced form test.

Table 6: Robustness of the structural form results (not for publication)

	Specification				
	(1)	(2)	(3)	(4)	(5)
2 RHS variables					
First stage partial F	16.83	16.83	19.18	8.5	12.14
Anderson p value	0.03	0.031	0.039	0.181	0.04
Overidentification p value	0.198	0.211	0.189	0.13	0.203
3 RHS variables					
Recent default point estimate	0.029	0.028	0.029	0.032	0.03
(p value)	0	0	0	0	0
First stage partial F	9.18	9.18	25.45	2.98	5.37
Anderson p value	0.016	0.016	0.016	0.228	0.02
Overidentification p value	0.32	0.309	0.192	0.182	0.298
4 RHS variables					
LIBOR point estimate	-0.094	-0.094	-0.094	-0.097	-0.095
(p value)	0	0	0	0	0
Future default point estimate	0.352	0.357	0.341	0.243	0.365
(p value)	0.029	0.029	0.027	0.116	0.03
Reserves point estimate	-0.702	-0.703	-0.693	-0.707	-0.695
(p value)	0.006	0.007	0.005	0.002	0.006
Recent default point estimate	0.015	0.015	0.016	0.019	0.017
(p value)	0.004	0.006	0.002	0.024	0.003
First stage partial F	6.76	6.76	19.43	2.88	4.09
Anderson p value	0.012	0.012	0.009	0.208	0.015
Overidentification p value	0.46	0.46	0.38	0.349	0.478

Column 1 is our benchmark specification. Column 2 uses only quantities as weights fir the spreads. Column 3 used quadratic discounting for recent default. Column 4 uses a future default measure where future arrears are split evenly among past contracts. Column 5 includes a 3 year grace period in future default. The sample size is 149 in column 1-4, and 150 in column 5.