Graz Economics Papers – GEP

[1]

GEP 2012–03

# The Ideal Loan and the Patterns of Cross-Border Bank Lending

Bettina Brüggemann, Jörn Kleinert, Esteban Prieto

September 2012

Department of Economics
Department of Public Economics
University of Graz

# The Ideal Loan and the Patterns of Cross-Border Bank Lending


### Abstract

A typical loan offer is a differentiated product with various negotiated characteristics (maturity, amount, timing, collateral, disclosure requirements) which involve costs that go beyond the mere interest rate. Taking into account all costs, a firm chooses the cost minimizing loan offer. Based on this decision criterion, we derive the probability of a firm from country $i$ to choose a loan contract from a bank in country $j$. We use this probability to derive a gravity equation for cross-border bank loans. Finally, we estimate the gravity equation based on the theoretical model controlling for the unobserved heterogeneity proposed by our theoretical model.



**Keywords and Phrases:** Product differentiation, Gravity Equation, Cross-border bank lending

**JEL Classification Numbers:** L14, F34, G21



Bettina Brüggemann
Department of Money and Macroeconomics
University of Frankfurt
Grüneburgplatz 1
D-60323 Frankfurt/Main, Germany

bettina.brueggemann@hof.uni-frankfurt.de

Jörn Kleinert
Department of Economics
University of Graz
Universitätstraße 15/F4
A-8010 Graz, Austria

joern.kleinert@uni-graz.at

Esteban Prieto
Department of Economics
University of Tübingen
Mohlstraße 36 D-72074 Tübingen, Germany

esteban.prieto@wiwi.uni-tuebingen.de

# 1 Introduction

International bank lending can be explained surprisingly well by gravity forces that also explain international trade flows so well (see among others Buch 2005). The success of these gravity forces is puzzling given that transport costs, which are often used to explain the negative effect of distance on trade, should not be important for cross-border loans. While the gravity model has often been applied in the context of international finance in general and banking activities in particular, there are only very few attempts to explain the underlying theoretical source of the success of the gravity equation in explaining cross-border financing activities (Martin & Rey 2004, Okawa & van Wincoop 2012). Moreover, the existing approaches focus on explaining cross-border equity positions. We, in contrast, provide a theoretical foundation for a gravity equation in cross-border bank lending. Our goal is (i) to study the underlying economic relationships which allow for a meaningful interpretation of the estimated coefficients and (ii) to search for the proper specification of the gravity equation. Our theory gives an explanation for the success of the gravity equation in explaining cross-border bank lending data. We pay particular attention to the role of distance, because the negative effect of distance in the transfer of "weightless" loans is often regarded as puzzling. We interpret the importance of the distance between the firm and the envisaged banking market as being due to the fact that distance raises the firm's costs when it searches for the best lender of funds for a particular project. In addition, distance also increases the bank's costs when monitoring the firm.

Our theoretical model starts from the observation that a loan offer has various crucial characteristics in addition to the interest rate (e.g. maturity, amount, timing, collateral, disclosure requirements).[2] Comparing two offers is therefore not trivial and the "ideal" offer does not necessarily feature the lowest interest rate. Moreover, a loan contract involves two parties and both must agree on the deal. Since in our study we rely on aggregate data and cannot observe all the relevant characteristics of the loan offers in question, we include a stochastic (i.e. unobserved) component to account for unobservable relationship-specific elements. Firms choose their preferred offer depending on both the observable part and the components that are unobservable to the researcher. We assume that these unobserved parts are independent and identically distributed. Assuming further that the cost minima are Gumbel extreme value distributed, the probability that a firm chooses the offer made by a particular bank can be derived. Summing over all banks from a particular country and the total need for external finance in the country where the firm is located, we compute total cross-border loans between the lending and the borrowing country. The resulting equation bears a strong resemblance to a gravity equation for bank loans.

After having derived this gravity equation, we estimate it using confidential locational cross-border bank lending data from the *Bank of International Settlement*. This data allows for a direct empirical test of our theory. The estimation results support the predictions of the theoretical model. We confirm that

[2] See also Francis et al. (2009) for empirical evidence that, in addition to loan rates (or loan spreads), bank loan contract also differ with respect to non-price components such as covenants and collateral.

cross-border bank lending decreases with the distance between two countries. The negative effect of the distance coefficient survives various robustness tests. It even prevails after controlling for determinants which capture the ability of country pairs to reduce contracting, search and monitoring costs, and the quality and effectiveness of the lending country's banking system.

The gravity equation in economics has traditionally been used in the context of international trade, but has also been applied to other economically relevant cross-border activities. As mentioned above, there are only few attempts to find a theoretical justification for the use of the gravity equation in finance. Martin & Rey (2004) use a general equilibrium framework to model assets as imperfect substitutes that insure against different risks and account for transaction costs in cross-border asset trade. Portes & Rey (2005) use the structural gravity equation from Martin & Rey (2004) to explain cross-border asset holdings. Okawa & van Wincoop (2012) use a portfolio approach to derive a gravity equation for equity. While resulting in a very similar equation, their approach is very different from ours. They start from a representative investor holding a portfolio of different assets. In contrast to their approach, we focus on cross-border bank loans. Central to our analysis are international debt markets and not cross-border asset markets, as we derive a gravity equation based on the analysis of the relationship between a firm demanding a bank loan and the potentially lending bank.

Markets for bank loans are decentralized without a unique market-clearing price. In that they differ from equity or bond markets where the assumption of a market clearing price can be justified much easier. Moreover, the size of cross-border loans' markets is large enough to warrant a distinct analysis. World-wide debt markets are at least as large as assets markets and a large share of debts falls on loans. Usually, loans are given to a customer by a bank after an auditing of project and firm. This, however, requires much more specific knowledge and time than deciding about an equity purchase because firm and bank are locked in a one-to-one relationship after closing the loans contract. Equities are usually easily tradable at low costs and at any point in time while loans are not. Yet, loans' repayments are by far less volatile than equity returns. Both arguments taken together imply that the portfolio diversification motive does not weigh as heavily for loans as it does for equity. Risks must be dealt with through the gathering of information and a larger degree of involvement in the activities of the firms. These characteristics of the loan market call for a partial equilibrium approach in explaining the loans pattern. This implies in particular that we analyze the firm-bank relationship concerning a particular loan in isolation, i.e. without optimizing over the whole portfolio of loans a bank holds and the whole portfolio of banks that a firm is indebted to.

Our paper is most closely related to Head & Ries (2008), who build on Anderson et al. (1992) to derive a gravity representation for FDI. FDI is there seen as an outcome of the market for corporate control. Head & Ries (2008) model a matching between two firms in an auction for the specific assets of the acquisition target. The firm with the highest ability to use the assets profitably can afford to pay the highest price and therefore controls the assets. We have a similar auction in mind, when we model a loan as outcome of a firms search for the cheapest supplier of bank credit. Hence, instead of assuming a representative bank with several loans, we are closer to the “ideal loan approach”.

# 2 A theory of aggregate cross-border bank loans

We start with a single firm that searches for a loan that provides the best possible conditions among different banks. These loans differ in their characteristics, including the interest rate, fixed fees, collateral, reporting requirements and other details that such a contract might feature. The OECD for instance states in the description of their interest rates in their Financial Indicator (MEI): "...rates vary not only because of inflation ... but also because of a number of other influences, including the amount, purpose and period of the transaction, the credit-worthiness of the borrower, the collateral offered and/or guarantees/guarantors available, the competition for the transaction, government policy. As a consequence, there will be numerous rates applying to a large number of transactions that are in effect at any one time in any one country." (OECD (2011)).

Some of these features are observable to the researcher, others are not. Using the observable part, we predict the probability that a particular bank $k$ from a foreign country $j$ provides a loan to a representative firm $g$ in country $i$. Using only country-level variables, this probability is not overly informative with respect to one individual bank. It can however be used to calculate the probability that a firm in country $i$ chooses a loan from country $j$, no matter what bank it chooses in particular. We then use this probability to compute the expected volume of loans that banks from country $j$ give to firms in $i$.

## 2.1 Firm-bank relationship

Assume that a firm $g$ from country $i$ searches for a bank loan in $N$ relevant countries including its home country. This firm chooses bank $k$ in country $j$ if this very bank offers the loan for the lowest overall cost $c_{igjk}$. The costs incurred by the borrowing firm $g$ are specific to its loan contract with bank $k$. They depend on the interest rate $r_{igjk}$ for the loan on which the firm and the bank have agreed, but also on other factors $a_{igjk}$ that capture loan- and bank-specific cost components. We might think of these additional factors captured by $a_{igjk}$ as contract-specific features such as the amount the firm borrows, the collateral, or non-monetary variables like timing and maturity. Other factors included in $a_{igjk}$ are unspecific to the contract, e.g. they refer to the efficiency of the bank involved. We assume that all these characteristics contained in $a_{igjk}$ and the interest rate $r_{igjk}$ can be collapsed in one variable, overall costs $c_{igjk}$. These costs are used to compare any two loan offers.

Although, in our view, it is up to the firm to approach banks and ask for loan offers, banks are by no means passive or non-optimizing here. We think of the whole procedure as consisting of three steps: First, firms ask a bank for the conditions they can offer in a loan for a specific project. Second, the firm and the bank negotiate the conditions and the bank makes a take-it-or-leave-it offer. The firm does or does not accept the offer depending on the conditions offered by the other banks it had asked for a loan. Hence, the offered conditions depend not only on the characteristics of the bank such as its evaluation of the firm and the project, but also on its alternative lending opportunities.

Usually, the exact amount of overall costs $c_{igjk}$ are unobservable to the researcher. However, some elements of the costs are observable. The interest rate $r_{igjk}$ on which the two partners agree is certainly affected by the prevailing interest rate $r_j$ in the country where the bank is located. We think of the prevailing interest rate in $j$, $r_j$, as the average of the interest rates of all loan contracts between firms and banks in this country. The contracting costs and the costs of monitoring $t_{igjk}$ incurred by the bank also affect this interest rate. Although the relationship-specific monitoring and contracting costs $t_{igjk}$ are not observable either, they depend systematically on observable variables such as distance between the bank and the firm and other distance-related variables. This observable part of the monitoring costs can have several dimensions, such as physical and cultural distance, familiarity with the business, institutional features in the debtor country or the specific financing needs. The variable $\tau_{ij}$ denotes the part of the monitoring cost $t_{igjk}$ that is determined by the average distance between the two countries involved and other distance-related variables. We can now decompose he interest rate that lender and borrower agree on, $r_{igjk}$, into an observable part and an unobservable part $\varepsilon$ with zero mean:

$$r_{igjk} = r_j + \tau_{ij} + \varepsilon_{igjk}.$$

Using the same logic, we rewrite the overall costs $c_{igjk}$ as a function of observable variables at the country level and firm-bank-relationship specific unobservable parts that are collapsed in the error term $\varepsilon_{igjk}$ which has zero mean. The borrowing firm's costs of a loan from bank $k$ can then be described as

$$c_{igjk} = r_j + \tau_{ij} + a_j + \varepsilon_{igjk},$$

where $a_j$ stands for average bank characteristics in country $j$ that might affect the cost of the loan. It can be thought of as an average of the relevant loan- and bank-specific cost factors $a_{igjk}$. Deviations from this average are random and therefore also captured by the error term. Note that our analysis does not depend on the assumption of a representative firm. The model is flexible enough to take into account any number of firms $n$ in country $i$. Moreover, since the observable part of the cost function does not depend on firm-specific factors, we can interpret this observable part as the average costs of borrowing from a foreign bank $k$ in country $j$.[3] We will stick to this interpretation of the above cost function from now on.

Average costs, $\bar{c}_{ij} = r_j + \tau_{ij} + a_j$, include only country averages and bilateral distance-related variables. The bank-specific deviations from these country-specific variables are random from the perspective of the researcher, with $\varepsilon_{igjk}$ collecting all random elements. We can then write the cost function as

[3] Firm heterogeneity only enters the cost function through the error term $\varepsilon_{igjk}$, which has zero mean by assumption. When averaging over all firms it would therefore simply cancel out, which renders our model very flexible with regard to the number of firms included.

$c_{igjk} = \bar{c}_{ij} + \varepsilon_{igjk}$. Firm $g$ from country $i$ chooses bank $k$ from country $j$ to take out a loan if this loan minimizes its costs. The probability that firm $g$ from country $i$ chooses bank $k$ from country $j$ is then given by

$$\begin{aligned}
\mathbf{P}_{igjk} &= \Pr\left(c_{igjk} = \min\{c_{l1}\cdots c_{ln_l}; l = 1\cdots N\}\right) \\
&= \Pr\left(\bar{c}_{ij} + \varepsilon_{igjk} < \bar{c}_{il} + \varepsilon_{igl1}; \cdots; \bar{c}_{ij} + \varepsilon_{igjk} < \bar{c}_{il} + \varepsilon_{igln_l};\ \ l = 1\cdots N, jk \neq lh\right) \\
&= \Pr\left(\bar{c}_{ij} - \bar{c}_{il} + \varepsilon_{igjk} < \varepsilon_{igl1}; \cdots; \bar{c}_{ij} - \bar{c}_{il} + \varepsilon_{igjk} < \varepsilon_{igln_l};\ \ l = 1\cdots N, jk \neq lh\right) \\
&= 1 - \Pr\left(\bar{c}_{ij} - \bar{c}_{il} + \varepsilon_{igjk} \geq \varepsilon_{igl1}; \cdots; \bar{c}_{ij} - \bar{c}_{il} + \varepsilon_{igjk} \geq \varepsilon_{igln_l};\ \ l = 1\cdots N, jk \neq lh\right) \\
&= \underbrace{\prod_{l=1}^{N}\prod_{h=1}^{n_l}}_{lh\neq jk} \left[1 - F(\bar{c}_{ij} - \bar{c}_{il} + x)\right], l = 1\cdots N,
\end{aligned}$$

where $l = 1\cdots N$ is the country index and the index $h = 1\cdots n_l$ labels the different banks in country $l$ that are considered by a firm in country $i$. The random term $\varepsilon$ is i.i.d. Its cumulative distribution function is denoted by $F$, the corresponding density $f$. For any given realization $x$ of $\varepsilon_{igjk}$, bank loan from $k$ in country $j$ are chosen with probability density

$$\underbrace{\prod_{l=1}^{N}\prod_{h=1}^{n_l}}_{lh\neq jk} \left[1 - F(\bar{c}_{ij} - \bar{c}_{il} + x)\right],$$

where we multiply over all eligible bank loan variants $h$ in all countries $l$. None of the deterministic variables in this equation varies from bank to bank; the variation is only between the countries. We can therefore rewrite it as $\prod_{l=1}^{N}\left[1 - F(\bar{c}_{ij} - \bar{c}_{il} + x)\right]^{n_l}$, keeping in mind that $n_j$ excludes the firm-bank relationship in question, $igjk$. Accounting for all possible realizations of $x$, the probability $\mathbf{P}_{igjk}$ of choosing bank $k$ in country $j$ can be calculated using:

$$\mathbf{P}_{igjk} = \int_{-\infty}^{\infty} f(x) \prod_{l=1}^{N} \left[1 - F(\bar{c}_{ij} - \bar{c}_{il} + x)\right]^{n_l} dx. \tag{1}$$

The probability that a firm $g$ from country $i$ chooses a bank $k$ from country $j$ depends on the average cost characteristics in country $j$, on the bilateral characteristics, and on the number of banks approached in country $j$, $n_j$.

## 2.2 Cross-border bank loans at the country level

$[1-F(\bar{c}_{ij}-\bar{c}_{il}+x)]^{n_l}$ is the distribution of the minima in country $l$, where $F$ is unknown. For a large number of banks $n_l$ however we can work with an asymptotic distribution which does not depend on the sampled distribution function $F$. For a number of families of distributions like normal, log-normal, logistic, and exponential distributions, this limiting distribution is the Gumbel distribution. The Gumbel distribution has a double exponential form. For minimum values of residuals with zero mean, it is given by

$$1-G(x)=\exp\left[-\exp\left(\frac{x-\mu}{\sigma}\right)\right] \tag{2}$$

where $\sigma$ is the constant scale parameter describing the "horizontal stretching" of the distribution, and $\mu$ is the location parameter of the Gumbel distribution. The location parameter $\mu$ is given by $\mu=\sigma_\varepsilon \ln n_l$ with $\sigma_\varepsilon$ denoting the standard deviation of the residuals of the underlying normal distribution and $n_l$ denoting the number of banks in country $l$ that could have been approached by firm $g$. The scale parameter $\sigma$ equals the standard deviation $\sigma_\varepsilon$ of the normally distributed residuals $\varepsilon$.[4] The minima in country $l$ are therefore distributed according to

$$1-G(\bar{c}_{ij}-\bar{c}_{il}+x)=\exp\left[-\exp\left(\frac{\bar{c}_{ij}-\bar{c}_{il}+x-\sigma\ln n_l}{\sigma}\right)\right] \tag{3}$$

The minima in each country $l$ are Gumbel distributed according to (3). We are interested in the minimum values over all countries and therefore in the minimum of the minima. From the cumulative distribution function $G(x)$, the density function $g(x)$ can be derived as

$$g(x)=\frac{1}{\sigma}\exp\left(\frac{x-\mu}{\sigma}\right)\left\{\exp\left[-\exp\left(\frac{x-\mu}{\sigma}\right)\right]\right\}$$

Using the limiting distribution, the probability that firm $g$ from country $i$ chooses a loan provided by bank $k$ from country $j$ is thus given by

$$\mathbf{P}_{igjk}=\int_{-\infty}^{\infty}\frac{1}{\sigma}\exp\left(\frac{x-\mu}{\sigma}\right)\left\{\exp\left[-\exp\left(\frac{x-\mu}{\sigma}\right)\right]\right\} \prod_{l\neq j}^{N}\exp\left[-\exp\left(\frac{\bar{c}_{ij}-\bar{c}_{il}+\sigma\ln n_l+x-\mu}{\sigma}\right)\right]dx. \tag{4}$$

Solving this integral (see Appendix) yields the following expression for the probability of a firm $g$ in $i$ choosing bank $k$ from $j$:

$$\mathbf{P}_{igjk}=\frac{\exp\left(-\frac{\bar{c}_{ij}}{\sigma}\right)}{\sum_{l=1}^{N}n_l\exp\left(-\frac{\bar{c}_{il}}{\sigma}\right)}. \tag{5}$$

[4] See Mood et al. (1974): 258-260.

This equation gives us the same probability for every bank $k$ from country $j$ that a firm from country $i$ takes into consideration for borrowing. As we use only aggregate data, bank-level probabilities are not very informative. We are much more interested in the probability that a firm in country $i$ chooses any one of the banks from country $j$ to take out a loan. We can obtain an expression for this probability by simply aggregating the probabilities over all $n_j$ eligible banks in country $j$. The probability of a firm choosing any bank in country $j$ can then be written as

$$\mathbf{P}_{igj} = \frac{n_j \exp\left(-\frac{\bar{c}_{ij}}{\sigma}\right)}{\sum_{l=1}^{N} n_l \exp\left(-\frac{\bar{c}_{il}}{\sigma}\right)}. \tag{6}$$

The number of eligible banks does not have to be equal to the total number of banks in a country. It is more likely that a firm will only screen some of the potential lenders. This number of banks screened depends on the distance between country $i$ and country $j$, and on other characteristics of the banking sector. These factors enter linearly and are therefore already accounted for by $\tau_{ij}$ and $a_j$.

To arrive at an expression for the expected total amount of bank loans floating from country $j$ to country $i$, $BA_{ji}$, we simply multiply the above probability by the total amount of bank loans in country $i$, $BL_i$:

$$BA_{ij} = \frac{n_j \exp\left(-\frac{r_j+\tau_{ij}+a_j}{\sigma}\right)}{\sum_{l=1}^{N} n_l \exp\left(-\frac{r_l+\tau_{il}+a_l}{\sigma}\right)} BL_i. \tag{7}$$

In this equation, we have also replaced the country-specific cost variables by their different components which we had defined in the beginning of this section. Total loans are affected by the prevailing interest rate in country $j$, $r_j$, by the monitoring costs $\tau_{ij}$ that a bank in country $j$ has to incur when lending to a firm in country $i$, and the average quality of country $j$'s banks, captured by the average bank characteristics $a_j$. The same effects, just summed up for all countries, can be found in the sum in the denominator, which describes the relationship between country $i$ and the rest of the countries that were taken into consideration during the search for a bank loan. This sum discounts the influence of the country-$j$-variables in the numerator.

## 2.3 Comparison to gravity in equity

Before we turn to estimating the derived gravity equation, we pause to compare our approach to two other attempts that aim at giving the gravity equation a theoretical foundation. The main difference to our approach is that we look particularly at the cross-border loans segment of banking activities. To do that, we model a firm-bank relationship, where the firm takes a more active part but the bank is also optimizing: the firm approaches different banks to obtain a loan offer and compares the different loan offers it receives. We model the relationship like this because we believe that the loan demanding firm has this active role in the bank lending market. There are other segments of the financial markets such as the bond market and the equity market where the bank itself is searching for the optimal portfolio.

The equity segments in particular have been in the focus of research on cross-border bank activities. Although, in our setup, the bank might not decide about loans and equity (which we do not consider) independently, we do not have an omitted variable bias as the loan offer by the bank is worked out taking into account the possibility of alternative investment opportunities for the financial means of the bank. In this way, we can analyze a large number of firm-bank relationships without considering the deposit side explicitly.

Martin & Rey (2004) propose a two-country model with two periods to study portfolio (equity) investments. Individuals invest in differentiated assets in order to hedge against idiosyncratic risk in a stochastic world where each assets yields a payment in one out of $L$ stages of nature that can occur in period two. Individuals are risk-averse with constant relative risk aversion. Together with the structure of the payoffs in the second period, risk aversion creates a love of diversity similar to the love of variety in trade models which yields a CES-like structure of asset holdings. Based on this structure Martin & Rey (2004) derive a gravity equation which is in its structure very close to gravity equations from trade models but proposes different "mass" and "distance cost" variables. In addition, the expected return on assets in the investment country shows up in the equation. There are no fixed effects proposed which stems from the two-country structure. Coeurdacier & Martin (2009) extend the model to $n$ countries and yield a gravity equation with a multilateral resistance term for the home and the host country.

In contrast to Martin & Rey (2004), who basically derive a gravity equation for financial holdings when countries trade claims on Arrow Debreu securities, Okawa & van Wincoop (2012) derive their gravity equation from standard static portfolio theory. In their model, investors can hold claims on the risky asset from each of the many different countries. Each of these assets is affected by a global and a country-specific risk factor. Okawa & van Wincoop (2012) also allow for trade in a risk-free asset as well as an asset whose return depends only on global risk. In addition, they introduce financial frictions that are mainly due to information asymmetries. Assuming asset market clearing, they derive a gravity equation where economic masses are given by the supply of equity in one country and total equity holdings in the other countries. Distance enters the equation through a relative friction term that relates bilateral financial frictions to multilateral resistance terms. Okawa & van Wincoop (2012) introduce three different, easily applicable methods of estimating these bilateral frictions based on a number of observables, such as distance, common language, or the same legal system. Thus, although all three theoretical approaches motivate a gravity equation, the proper specifications and the interpretation of the included variables are rather different.

# 3 Data and Empirical Specification

Equation (7) is a gravity equation. The dependent variable, $\mathrm{BA}_{ij}$, is the total amount of bank loans in the debtor country $i$ from lending country $j$. We construct $\mathrm{BA}_{ij}$ using year-end data on cross-border bank claims from the unpublished locational bilateral banking data provided by the Bank of International

Settlement (BIS). The sample that we use covers the period 2003 to 2006. Descriptive statistics for all variables used can be found in Table 1.[5]

There are two variables accounting for economic masses: total loans in the loan-receiving country, $\mathrm{BL}_i$, and the number of banks in the loan sending country, $n_j$. In line with gravity theory, both mass variables enter the equation with a positive sign, and hence, have a positive effect on bilateral bank loans. Our theory predicts that the coefficients on these variables are one. While total loans in the host country are available for most countries, the number of foreign banks $n_j$ is harder to come by. The problem is that $n_j$ refers to all foreign banks over which a firm searches for the "ideal loan" which is necessarily unknown. We proxy $n_j$ by the overall size of the loan granting country measured by total assets of the banking system $\mathrm{TBA}_j$. However, the number of banks over which a firm searches depends also on other determinants like the intensity of contact between the two countries or the degree to which firms in country $i$ are familiar with the banks in country $j$. In the theoretical framework, this feature is captured by the dependence of the number of banks a firm screens on the distance between the firm and the bank. In the empirical framework, these characteristics are accounted for by the physical distance and other distance-related variables between the two countries summarized in $\tau_{ij}$. We take distance and distance-related variables from the *CEPII*.

The different cost components have a negative effect on the volume of cross-border bank loans. The bank-firm-specific cost function specified above, $c_{igjk} = r_j + \tau_{ij} + a_j + \varepsilon_{igjk}$ describes the costs incurred through the prevailing interest rate $r_j$, the monitoring costs $\tau_{ij}$, and the quality of the banking sector $a_j$. We proxy the monitoring costs by physical distance (in logs) and other distance-related variables. Note that in the theoretical setup, distance enters the equation at two different points: in the monitoring costs which strongly depend on the distance between the lending bank and the borrowing firm *and* in determining the actual number of banks per country $j$ that a firm considers for borrowing as described above. Since both effects enter linearly they can be added and estimated together. We approximate the loan originating country's lending rate $r_j$ by the average bank lending rate in country $j$.

The sum in the denominator in equation (7) is the same for every bilateral relationship of a particular receiving country $i$. This sum differs between any two receiving countries. In this sense the denominator is closely related to what in the trade literature has become known as multilateral (price) resistance terms (Anderson & van Wincoop 2003). In order to account for these "multilateral (cost) resistance terms" we introduce country $i$-(receiving country) specific fixed effects $D_i$. Once we introduce receiving-country dummy variables, the effect of the mass variable $BL_i$ is no longer identified. In the following, we therefore ignore the loan receiving-country mass variable keeping in mind that its effect is subsumed in the receiving-country fixed effects. The importance of receiving-country fixed effects becomes clear when looking at equation (7). This equation shows that neglecting to account for the sum in the denominator when estimating a gravity equation for bilateral bank lending introduces an omitted variable bias. The results would be biased because the multilateral resistance terms are correlated with $r_j$, $\tau_{ij}$ and $a_j$.

[5] See the data appendix for a detailed description.

Applying these considerations to equation (7) and adding an error term, we can express the empirical analogue to the gravity equation for cross-border bank lending as:

$$\mathrm{BA}_{ij} = \exp\left[(\beta_1 r_j + \beta_2 \tau_{ij} + \beta_3 a_j - D_i)\right] \mathrm{TBA}_j^{\beta_4} \varepsilon_{ij}. \tag{8}$$

The common procedure for estimating a gravity equation as given in equation (8) is to take logs of both sides, which yields

$$\ln \mathrm{BA}_{ij} = \beta_1 r_j + \beta_2 \tau_{ij} + \beta_3 a_j + \beta_4 \ln \mathrm{TBA}_j + D_i + \ln \varepsilon_{ij} \tag{9}$$

The parameters in equation (9) can be estimated using linear regression with country $i$-(receiving country) specific-fixed effects ($D_i$) to account for multilateral (cost) resistances. Fixed-effects ordinary least squares (FE-OLS) estimation produces consistent parameters estimates given that the variance of the error term $\varepsilon_{ij}$ in equation (8) is independent of the explanatory variables. If this condition is violated, linear least squares estimates are inconsistent. To cope with this possible inconsistency of the FE-OLS estimator, Silva & Tenreyro (2006) propose using a Poisson pseudo-maximum likelihood estimator which is robust to different patterns of heteroskedasticity. Poisson estimation allows estimating the gravity equation directly from its multiplicative form. The gravity equation in multiplicative from can be written as:

$$\mathrm{BA}_{ji} = \exp\left[\beta_1 r_j + \beta_2 \tau_{ij} + \beta_3 \ln a_j + \beta_4 \mathrm{TBA}_j + D_i\right] \varepsilon_{ij}. \tag{10}$$

This specification follows directly from the multiplicative form in equation (8).

According to the simulation results in Silva & Tenreyro (2006) and Westerlund & Wilhelmsson (2011), the bias introduced by log-linearizing a gravity equation can be substantial. In order to get a feeling about the importance of this issue in the context of international bank lending, we compare the results of both estimation procedures.

Another issue is that our theoretical model is derived in a static framework. In its standard form, the gravity equation deduced from the theoretical framework is therefore suitable for a cross-sectional estimation. However, it is straight forward to extent the empirical model to a panel analysis. Panel data and the related panel econometric methods have some well known advantages over cross-sectional econometric methods.[6] We therefore want to exploit the panel nature of our dataset. But we have to keep in mind, that the economic content of the parameter estimates from a panel estimation is different from their cross-sectional counterpart. The outcome from an estimation of a cross-section represents long-term relationships, whereas the parameter estimates of a panel regression might better be interpreted as short-term, within-group effects (Egger & Pfaffermayr 2003). Both interpretations go along with our theoretical model. Therefore, when we apply our gravity equation to panel data, the general setup re-

[6] See Baltagi 2001 for a discussion of this issue.

mains unchanged, but we need to take into account the time dimension in the sums and in the unobserved characteristics of the banking sector in country $j$ in order to get the structural relationships between the variables. The correct empirical specification in a panel setup thus reads as follows:

$$\ln \mathrm{BA}_{jit} = \beta_1 r_{jt} + \beta_2 \tau_{ij} + \beta_3 a_{jt} + \beta_4 \ln \mathrm{TBA}_{jt} + D_{it} + \ln \varepsilon_{ijt}. \tag{11}$$

The analogue multiplicative panel-data econometric specification is given by

$$\mathrm{BA}_{jit} = exp\left[\beta_1 r_{jt} + \beta_2 \tau_{ij} + \beta_3 a_{jt} + \beta_4 \ln \mathrm{TBA}_{jt} + D_{it}\right] \varepsilon_{ijt}. \tag{12}$$

In a panel setting, the multilateral resistance variables are time varying. Consequently, ignoring the time variation in the fixed effects can produce biased estimates because of an omitted variables bias. We therefore apply year-country fixed effects (i.e. time-varying country-specific fixed effects) for the loan receiving countries in our empirical specification.

The gravity representation is very flexible in including additional variables, because the explanatory variables specify monitoring costs $t_{igjk}$ and the effective number of banks $n_j$, both depending on distance, and the quality $a_j$ of the lending country's banking sector. Buch (2002) for instance uses (geographical) Distance, Language, EU membership, Capital controls and the Legal system to catch the different dimensions of the cost of a contract between lender and borrower from two different countries. In her empirical approach, the quality of the banking sector is reflected by GDP per capita, the Index of human development, and Restrictions in banking and finance. Papaioannou (2009) in contrast concentrates on political and institutional factors which he shows are also important determinants in explaining international bank lending.

The ambiguity of the gravity representation for international bank lending with respect to the variables to condition on complicates the estimation of the gravity equation. The quality of the banking sector and the effective number of banks in a country are largely unobserved. If the empirical choice on how to approximate these two quantities neglects one or more dimensions of the quality of the banking sector, the results of the estimation would clearly be biased due to omitted variables. To deal with this problem stemming from uncontrolled characteristics on the loan-originating country we include, on top of the loan-receiving-country fixed effects, a full set of (time-varying) fixed effects for the loan-originating country. These fixed effects capture all loan-originating country's characteristics including the effects of the interest rate and the mass variable in the loan sending country.[7]

In order to test the predictions of our theory concerning the variables included in equation (8), we proceed with a two-step approach. In a first step, we regress cross-border bank loans on a full set of time-varying loan-receiving and -originating country dummies and a vector of distance-related variables.

[7] See Baier & Bergstrand (2007), Baldwin & Taglioni (2006), and Head & Ries (2008) for similar approaches to deal with country effects in a panel setup.

To be concrete, in the first step we estimate regressions of the form

$$\mathrm{BA}_{jit} = exp\left[\beta\tau_{ij} + D_{jt} + D_{it}\right]\varepsilon_{ijt} \tag{13}$$

$$\ln \mathrm{BA}_{jit} = \beta\tau_{ij} + D_{jt} + D_{it} + \ln \varepsilon_{ijt}. \tag{14}$$

In the second step, we use the estimated loan originating country fixed effects $\tilde{D}_{jt}$ as the dependent variable in a regression on variables affecting cross border bank lending according to our theory:

$$\tilde{D}_{jt} = \gamma_1 r_{jt} + \gamma_2 a_{jt} + \gamma_3 \ln \mathrm{TBA}_{jt} + \xi_{tj} \tag{15}$$

We estimate this equation with weighted ordinary least squares using the inverse of the estimated standard errors of the fixed effects as weights. In this way we can account for the heteroskedasticity in the estimated fixed effects.

**?** show in an application of the gravity equation to international trade that the export country fixed effects can be successfully explained by the market size variable and the outward multilateral resistance terms of the exporting country. This outward multilateral resistance term results from their general equilibrium approach with free market entry. In our approach, we refrain from modeling a general equilibrium model with endogenous banks' market entry to explain the pattern of cross-border banking. Thus, the loan originating country fixed effect is not affected by a multilateral resistance term in our approach. It is determined only by the variables given in 15.

# 4 Results

In Table 2, we present the results of estimating the first stage gravity equation. Column (1) in Table 2 shows the results using the Poisson estimator which estimates the gravity equation in its multiplicative form (see equation 13). Column (2) and column (3) present results using the conventional fixed effects OLS approach (see equation 14). In column (2), the dependent variable is the logarithm of the stock of cross-border bank loans while in column (3) the dependent variable is constructed as the logarithm of one plus the stock of cross-border bank lending ($\ln(1+\mathrm{BA}_{ij})$). The latter approach is a common procedure to deal with zeros in international trade or financial holding data when estimating log-linearized gravity equations.

The Poisson estimator (column (1)) produces a distance coefficient which is in line with our theoretical prior. The point estimate of the distance coefficient is $-0.26$ and significantly smaller than the common, a-theoretical prior of $-1$ (Baldwin & Taglioni 2006). According to theory, firms tend to have more problems with appropriately screening remote foreign banking markets. The negative distance effect might also reflect that monitoring costs incurred by the lending bank increase with distance.

In addition to the geographical distance we include a number of variables related to contracting, search and monitoring costs. Again, this cost component has various dimensions. In order to better understand what lies behind the various costs of searching, monitoring and contracting we augment the specification with dummy variables indicating the presence of a common border (contiguity), a common official language (common language), a common legal origin (common legal origin), an active free trade agreement between the two countries (regional trade agreement) or a common currency (common currency).[8]

First, speaking the same language might reduce the cost of writing loan contracts. This should be related to an overall reduction in the transaction costs. Also, a common language should ex ante increase the probability that a firm applies to a particular bank. Both effects should lead to a larger aggregate lending volume between two countries. Second, two countries which share a common border generally *trade* relatively more. The border dummy might thus control for the fact that center to center distances might overestimate the effective distance between neighboring countries. According to the results in Table 2, two countries with the same official language do indeed have relatively higher bilateral cross border loan transactions, although its coefficient is only significant for the OLS estimation. A common border in contrast does not seem to have an independent effect on the volume of international bank lending.

Third, countries which are part of the same regional trade agreement or share a common currency also seem to have a higher volume of cross-border bank lending. The coefficients on both variables are positive and highly significant. The dummy variable indicating whether a country pair has the same legal origin exerts a positive and significant effect only when using the Poisson estimator. The same legal origin proxies a similar legal system or a comparable business conduct. Insofar as a common legal system or the similarity of the business conduct is able to reduce the time and resources devoted to writing a loan contract, transaction costs between these countries are smaller. Smaller transaction costs in turn favor cross-border lending between these country pairs.

Comparing the results obtained by using the Poisson estimator to the results by using the common FE-OLS estimator, we can observe that the absolute value of the distance coefficient in both OLS specifications is significantly larger than the corresponding coefficient produced by the Poisson estimator (point estimates of $-0.73$ and $-0.83$ compared to $-0.26$ for the Poisson estimator). This result resembles the finding of Silva & Tenreyro (2006) for gravity specifications for trade data. In fact, the OLS estimators seem to attribute a much larger effect of distance variables on the volume of bilateral lending compared to the Poisson estimator. A general picture which emerges from our result is that the point estimates produced by the two OLS models are very similar. This suggests that differences between the Poisson and

[8]The common legal origin variable is taken from La Porta et al. (1998), the regional trade dummy and the common currency dummy are from the CEPII Gravity Dataset available from the CEPII homepage, the remaining distance related variables are from the CEPII distance database.

OLS estimator are not due to truncation (because of zeros in the data) but rather due to log-linearization of the model in the presence of heteroskedasticity which leads to inconsistent estimates.

***Tracing the effects of source banking market characteristics***

In the second stage of the estimation procedure, we regress the estimated lending country fixed effects from the Poisson estimation on a variable approximating the number of banks $n_j$, on the interest rate in the lending country, $r_j$, and on variables approximating the quality of the lending countries banking system $a_j$ (see equation 7). We approximate $n_j$ by the overall size of the banking system in the loan granting country which we measure by total assets in the banking system. According to our theory we expect the size of the banking system to exert a positive effect on the volume of cross border bank loans. In the second stage, the mass variable should therefore enter the regression equation with a positive sign. The bank lending rate in the loan originating country represents a cost component in the gravity equation (7). Higher interest rates are associated with higher costs for a borrower when taking out the loan. Countries with higher average bank lending rates should therefore extend relatively less cross border bank assets and the interest rate should enter with a negative sign in the second stage regression.

In Table 3, we present the results from the regression using the estimated time-varying fixed effects as dependent variable. In column (1), we include as explanatory variables only the size of the banking market together with the interest rate in the lending country. The results of this baseline regression show that the size of the lending country's banking market and the average lending rate in this country both have the expected sign. Consistent with the predictions of our theoretical model the mass variable is positive and highly significant indicating that larger banking markets have relatively higher volumes of cross border bank lending. Also, countries which feature higher average bank lending rates seem to be the origin of lower volumes of international bank loans. This suggests that pure (interest rate) cost considerations play an important role for firms when accepting a loan offer. Hence, higher lending rates in the loan granting country act as a barrier to cross-border lending. Finally, the $R^2$ statistic reported at the bottom of Table 3 shows that the size of the banking system together with the bank lending rate explain the large bulk of variation in the estimated fixed effects. Only about 14% of the variation is left unexplained.

According to our theoretical model, the quality of the foreign banking market from which firms borrow is an important determinant of overall credit costs. In order to understand how characteristics of the banking system affect international bank lending, we add variables which characterize the profitability and the efficiency of the lending country's banking market in columns (2) to (7). As such these covariates might be able to approximate the term $a_j$ in equation (7). Specifically, we add the following variables to the baseline regression, one at a time: the 3 Bank Concentration Ratio, the Return on Assets, the Cost to Income Ratio, the Z-Score, the ratio of Total Bank Assets to GDP and the ratio of Total Bank Credit to

GDP.[9]

The 3 Bank Concentration ratio is a measure of the concentration in the banking system. The point estimate is positive indicating that more concentrated banking systems have higher volumes of cross border bank lending, but the estimate is not statistically significant. Return on Assets is an indicator for bank profitability. More profitable banks should be able to provide more competitive contract terms. We therefore expect the coefficient of this variable to be positive. The estimated coefficient has the wrong sign but is not significantly different from zero. The cost-income ratio, measured as the ratio of overhead cost to gross revenue, is a measure of non-operating cost efficiency. Higher values of the cost-income ratio indicate lower levels of cost efficiency. We expect that banking systems with lower ratios of overhead cost to revenues should be able to offer better loan terms which should lead to higher stocks of outstanding cross-border loans. The coefficient on the cost-income ratio has the expected negative sign but is not significantly different from zero. Z-Score measures a banking system's riskiness with higher values indicating less risky banks. The Z-Score enters the regression with a negative sign but is insignificant.

The last two variables used to approximate the quality of the banking system are the ratios of total bank assets to GDP as well as the ratios of total bank credit to GDP. Different from the first set of variables these two ratios are more reduced form constructs and lack a specific structural interpretation. However, given their reduced form nature, these variables control for a large array of quality related characteristics of a country: they stand for the general development of the banking sector in the loan-originating country. Indeed, the results in Table 3 show that both, the ratio of total bank assets to GDP and the ratio of total bank credit to GDP, have a positive and highly significant coefficient. Concerning the effects of the interest rates and the size of the banking market, we observe that they retain their expected negative and positive effects, respectively, and remain highly significant in each of the different regression specifications.

One caveat in the second stage of our estimation procedure is the small number of observations on which we draw our inference. The number of data points we have available for each year in our second step is limited by the number of countries reporting to the BIS (see the data appendix for details). In order to increase the sample size for the second step we perform as a robustness check the same set of regressions on a sample which pools all estimated fixed effects. The results, presented in Table **??**, do not change.

The last two variables used to approximate the quality of the banking system are the ratios of total bank assets to GDP as well as the ratios of total bank credit to GDP. Different from the first set of variables these two ratios are more reduced form constructs and lack a specific structural interpretation. However, given their reduced form nature, these variables control for a large array of quality related characteristics of a country: they stand for the general development of the banking sector in the loan-

[9] All variables intended to control for the quality of the banking system are taken from the Financial Structure Database collected by Beck et al. (2009).

originating country. Indeed, the results in Table 3 show that both, the ratio of total bank assets to GDP and the ratio of total bank credit to GDP, have a positive and highly significant coefficient. Concerning the effects of the interest rates and the size of the banking market, we observe that they retain their expected negative and positive effects, respectively, and remain highly significant in each of the different regression specifications.

***Accounting for bank lending rate endogeneity***

A valid objection to the regressions in Table 3 is that they might suffer from misspecification due to endogeneity. The reason is that the volume of international bank lending and the bank lending rate in a country are jointly determined in general equilibrium. Any approach to identify the causal effect of bank lending rates on (international) bank lending volumes without properly accounting for the simultaneous determination of these variables might therefore suffer from an endogeneity bias. Our strategy to deal with the endogeneity of bank lending rates in the regression is to instrument the interest rate with the corresponding central banks' official discount rate and the concentration ratio of the countries banking market. The central banks' official discount rate is set with respect to inflation and output gap and can therefore be expected to be not determined by the prime rate or bilateral cross-border bank lending. The concentration ratio of the banking market is certainly not determined by the contemporaneous market interest rate but very likely to affect it. Both variables do not affect bilateral cross-border bank lending beyond their effect on the prime rate in our sample. The lower part of Table 4 supports our choice of instruments.

Table 4 presents the results of the instrumental variables regression. We can observe that the instrumental variables approach leaves our main message unaltered. The coefficient on the size of the lending's countries banking system is still positive and highly significant, and the bank interest rate still exerts a negative effect and remains significant. The same holds for the variables controlling for the quality of the lending's country banking market: Return on Assets, the Cost to Income Ratio and the Z-score remain insignificant; the ratio of total bank assets to GDP and the ratio of total bank credit to GDP retain their positive effects and are still significant at the one percent level. Comparing the IV regression estimates with the corresponding OLS regression estimates shows that the instrumental variables approach produces point estimates which are consistently lower than the OLS point estimates. This indicates that the least squares estimates are systematically upward biased due the uncontrolled endogeneity of the bank lending rate.

For our instrumental variables approach to be a valid identification strategy, two conditions on the instruments need to hold: the central bank discount rates and the banking system concentration ratio should be uncorrelated with the volume of international bank lending, and both variables should be sufficiently correlated with the average bank lending rate in the country. Of course, a central bank's decision concerning the monetary policy strategy should not be affected by conditions related specifically to the international bank lending market. At the same time, the central bank discount rate is the main reference rate for short term funding for banks. The discount rate should therefore have a strong and systematic impact on the average bank lending rate in the country. The 3 Bank Concentration Ratio, as a measure of competition of the banking system, should be related to the pricing power of banks and therefore to the average level of the bank lending rate. Further, the fact that the the concentration ratio is insignificant in the OLS regression (see column (1) in table 3) indicates that there does not exists any significant correlation between international bank lending and the concentration ratio, thereby supporting our choice to exclude it from the main model.

The lower part of Table 4 presents formal tests on the validity of our instrumental variables setup. The first two test statistic correspond to the relevance of the instruments, that is whether the instrument variables are sufficiently correlated with the bank lending rate. We report the first stage $F$-statistic on the two instruments and the first stage (partial) $R^2$. The results show that the $F$-statistic is in most cases below the conventional critical threshold value of 10. A $F$-statistic below 10 is generally viewed as indication of weak instruments. We therefore use the Limited Information Maximum Likelihood (LIML) estimator which is most robust to weak instrument problems (Hahn et al. 2004). Furthermore, the partial $R^2$ of the first stage regression is very high with values well above 0.3. We therefore infer that the weakness of our instruments should not be a concern. The Hansen J-test concerns the validity of the instruments; that is, whether they are uncorrelated with the error term of the main equation. The null hypothesis is that the instruments can be excluded from the main equation. The result indicates that we can not reject the null hypothesis in any of the regressions. Finally, we conduct a formal test on the hypothesis that the bank lending rate should indeed be treated as endogenous. The null hypothesis of the Durbin-Wu-Hausmann test that the bank lending rate is exogenous can be rejected in the baseline specification and in the models featuring the the Return on Assets and the Cost to Income Ratio. In the last three models the test indicates that the bank lending rate can be safely treated as exogenous. This means that for these three cases the estimation and inference based on the standard OLS methods shown in Table 3 are superior to the instrumental variables results of Table 4.

# 5 Conclusion

The paper gives a theoretical foundation for the gravity equation in cross-border bank lending. Starting from the relationship between the borrowing firm and the lending bank, we derive the volume of bank credit between country $i$ and country $j$, based on a problem of cost minimization. The cost function not only consists of the direct credit costs represented by the interest rate but also depends on cost factors rooted in the specificity of the respective firms, banks and financial systems. The theoretical framework delivers a formulation for the amount of cross-border bank lending which resembles a gravity equation. The gravity equation of bank loans features multilateral (cost) resistance terms and unobserved lending-country characteristics which need to be accounted for when applying a gravity framework to bank lending data.

Using this theoretical gravity formulation, we then estimate a gravity equation for international bank lending. The results obtained from the empirical implementation of our structural gravity equation provide strong evidence in favor of the predictions derived from the theoretical model. We find that international bank lending is strongly affected by distance and distance-related variables such as common legal origin, a common currency or common membership in a free trade agreement. These findings give support to the importance of the non-interest rate costs in our theory, like the cost of setting up and enforcing a loan contract. Proxies for the effective number of banks and the interest rate in the loan originating country have effects as predicted by our theory. We also find that the general development of the banking system exerts a positive effect on bilateral bank lending.

## References

Anderson, J. E. & van Wincoop, E. (2003), 'Gravity with gravitas: A solution to the border puzzle', *American Economic Review* **93**(1).

Anderson, S., de Palma, A. & Thisse, J.-F. (1992), *Discrete Choice Theory of Product Differentiation*, MIT Press.

Baier, S. L. & Bergstrand, J. H. (2007), 'Do free trade agreements actually increase members' international trade?', *Journal of International Economics* **71**(1).

Baldwin, R. & Taglioni, D. (2006), Gravity for dummies and dummies for gravity equations, Cepr discussion papers, C.E.P.R. Discussion Papers.

Baltagi, B. (2001), *Econometric Analysis of Panel Data, $2^{nd}$ edition*, John Wiley & Sons.

Beck, T., Demirguc-Kunt, A. & Levine, R. (2009), Financial institutions and markets across countries and over time - data and analysis, Policy Research Working Paper Series 4943, The World Bank.

Buch, C. M. (2002), 'Are banks different? evidence from international data', *International Finance* **5**(1).

Buch, C. M. (2005), 'Distance and international banking', *Review of International Economics* **13**(4).

Coeurdacier, N. & Martin, P. (2009), 'The geography of asset trade and the euro: Insiders and outsiders', *Journal of the Japanese and International Economies* **23**(2).

Egger, P. & Pfaffermayr, M. (2003), 'The proper panel econometric specification of the gravity equation: A three-way model with bilateral interaction effects', *Empirical Economics* **28**(3).

Francis, B., Hasan, I., Koetter, M. & Wu, Q. (2009), The effectiveness of corporate boards: Evidence from bank loan contracting, CEI Working Paper Series 2009-08, Center for Economic Institutions.

Hahn, J., Hausman, J. & Kuersteiner, G. (2004), 'Estimation with weak instruments: Accuracy of higher-order bias and mse approximations', *Econometrics Journal* **7**(1), 272–306.

Head, K., Mayer, T. & Ries, J. (2010), 'The erosion of colonial trade linkages after independence', *Journal of International Economics* **81**(1), 1–14.

Head, K. & Ries, J. (2008), 'Fdi as an outcome of the market for corporate control: Theory and evidence', *Journal of International Economics* **74**(1).

La Porta, R., de Silanes, F. L., Shleifer, A. & Vishny, R. W. (1998), 'Law and finance', *Journal of Political Economy* **106**(6).

Martin, P. & Rey, H. (2004), 'Financial super-markets: size matters for asset trade', *Journal of International Economics* **64**(2).

Mood, A. M., Graybill, F. A. & Boes, D. (1974), *Introduction to the theory of statistics*, third edn, McGraw-Hill.

OECD (2011), Monthly monetary and financial statistics (mei), Technical report.

Okawa, Y. & van Wincoop, E. (2012), 'Gravity in international finance', *Journal of International Economics* .

Papaioannou, E. (2009), 'What drives international financial flows? politics, institutions and other determinants', *Journal of Development Economics* **88**(2).

Portes, R. & Rey, H. (2005), 'The determinants of cross-border equity flows', *Journal of International Economics* **65**(2).

Silva, J. M. C. S. & Tenreyro, S. (2006), 'The log of gravity', *The Review of Economics and Statistics* **88**(4).

Westerlund, J. & Wilhelmsson, F. (2011), 'Estimating the gravity model without gravity using panel data', *Applied Economics* **43**.

# 6 Tables

Table 1: Descriptive statistics

$BA_{ij}$ are total loans form country $j$ to country $i$ measured in million US-Dollars.

| Variable | $N$ | Mean | Std. Deviation | Min | Max |
|---|---|---|---|---|---|
| $BA_{ij}$ | 6331 | 3135 | 22741 | 0 | 945509 |
| distance | 6331 | 8.46 | 0.94 | 5.15 | 9.88 |
| contiguity | 6331 | 0.03 | 0.18 | 0.00 | 1.00 |
| common language | 6331 | 0.13 | 0.33 | 0.00 | 1.00 |
| common legal origin | 6331 | 0.31 | 0.46 | 0.00 | 1.00 |
| Regional Trade Agreement | 6331 | 0.29 | 0.45 | 0.00 | 1.00 |
| Common Currency | 6331 | 0.05 | 0.22 | 0.00 | 1.00 |
| Banking Market $\text{Size}_j$ | 6331 | 13.38 | 1.45 | 9.23 | 15.89 |
| Interest $\text{Rate}_j$ | 6331 | 7.08 | 9.42 | 1.66 | 67.08 |
| 3 Bank Concentration Ratio | 6331 | 0.63 | 0.21 | 0.23 | 0.96 |
| Return on Assets | 6331 | 0.01 | 0.01 | -0.01 | 0.05 |
| Cost Income Ratio | 6331 | 0.68 | 0.13 | 0.42 | 1.13 |
| Z-Score | 5852 | 8.97 | 5.38 | 2.69 | 44.74 |
| Bank Asset to $\text{GDP}_j$ | 6331 | 1.19 | 0.42 | 0.36 | 1.84 |
| Bank Credit to $\text{GDP}_j$ | 6331 | 1.03 | 0.44 | 0.13 | 1.75 |

## Table 2: Panel gravity equation for cross-border bank lending

The dependent variable $BA_{ij}$ are assets of reporting country $j$ in country $i$. PPML stands for Poisson pseudo maximum likelihood estimation. Fixed effects for both lending and borrowing countries account for all country-specific independent variables in all three estimations. Robust standard errors in brackets. **, * indicate significant at the 1 and 5% level.

| | PPML | OLS | OLS (1+$BA_{ij}$) |
|---|---|---|---|
| distance | -0.26** | -0.73** | -0.83** |
| | [0.086] | [0.065] | [0.059] |
| contiguity | -0.02 | -0.17 | -0.33* |
| | [0.171] | [0.174] | [0.179] |
| common language | 0.40 | 0.81** | 0.89** |
| | [0.205] | [0.103] | [0.100] |
| common legal origin | 0.22* | 0.13 | -0.03 |
| | [0.087] | [0.076] | [0.073] |
| Regional Trade Agreement | 0.39** | 0.69** | 0.66** |
| | [0.136] | [0.132] | [0.120] |
| Common Currency | 0.94** | 1.46** | 1.85** |
| | [0.124] | [0.125] | [0.129] |
| $N$ | 6331 | 4947 | 6331 |
| $R^2$ | 0.857 | 0.64 | 0.66 |

## Table 3: Second-stage pooled cross-section regression

The dependent variable is a vector of the estimated fixed effects from the first stage gravity specification. All year-specific loan originating country fixed effects have been pooled. Robust standard errors in brackets. **, * indicate significant at the 1 and 5% level.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Banking Market Size$_j$ | 1.14** | 1.18** | 1.13** | 1.16** | 1.07** | 1.08** | 1.13** |
| | [0.116] | [0.153] | [0.134] | [0.111] | [0.112] | [0.109] | [0.113] |
| Interest Rate$_j$ | -0.10** | -0.09** | -0.10** | -0.09** | -0.10** | -0.07** | -0.07** |
| | [0.010] | [0.012] | [0.010] | [0.010] | [0.008] | [0.008] | [0.009] |
| 3 Bank Concentration Ratio$_j$ | | 0.75 | | | | | |
| | | [0.775] | | | | | |
| Return on Assets$_j$ | | | -2.84 | | | | |
| | | | [20.494] | | | | |
| Cost Income Ratio$_j$ | | | | -1.44 | | | |
| | | | | [1.072] | | | |
| Z-Score$_j$ | | | | | -0.04 | | |
| | | | | | [0.026] | | |
| Bank Asset to GDP$_j$ | | | | | | 1.08** | |
| | | | | | | [0.220] | |
| Bank Credit to GDP$_j$ | | | | | | | 1.13** |
| | | | | | | | [0.233] |
| $N$ | 50 | 50 | 50 | 50 | 48 | 50 | 50 |
| $R^2$ | 0.859 | 0.863 | 0.859 | 0.864 | 0.857 | 0.892 | 0.897 |

Table 4: Second-stage pooled cross-section instrumental variables regression

The dependent variable is a vector of the estimated fixed effects from the first stage gravity specification. All year-specific loan originating country fixed effects have been pooled. The variable INTEREST RATE$_j$ is instrumented by the country's central bank discount rate and the 3 Bank Concentration Ratio. Robust standard errors in brackets. **, * indicate significant at the 1 and 5% level

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Banking Market Size$_j$ | 1.09** | 1.13** | 1.11** | 1.05** | 1.07** | 1.12** |
| | [0.119] | [0.126] | [0.110] | [0.111] | [0.110] | [0.121] |
| Interest Rate$_j$ | -0.12** | -0.14** | -0.12** | -0.11** | -0.08** | -0.08** |
| | [0.019] | [0.020] | [0.017] | [0.016] | [0.020] | [0.025] |
| Return on Assets$_j$ | | 21.6 | | | | |
| | | [23.899] | | | | |
| Cost Income Ratio$_j$ | | | -1.23 | | | |
| | | | [1.091] | | | |
| Z-Score$_j$ | | | | -0.05 | | |
| | | | | [0.027] | | |
| Bank Asset to GDP$_j$ | | | | | 1.00** | |
| | | | | | [0.300] | |
| Bank Credit to GDP$_j$ | | | | | | 1.03** |
| | | | | | | [0.360] |
| $N$ | 50 | 50 | 50 | 48 | 50 | 50 |
| $R^2$ | 0.847 | 0.839 | 0.854 | 0.854 | 0.891 | 0.896 |
| weak identification test | | | | | | |
| F-statistic | 8.86 | 8.42 | 8.56 | 16.41 | 5.25 | 5.65 |
| p-value | 0.001 | 0.001 | 0.001 | 0.000 | 0.009 | 0.007 |
| 1st Stage (partial) $R^2$ | 0.485 | 0.344 | 0.481 | 0.579 | 0.349 | 0.309 |
| overidentification test | | | | | | |
| Hansen J-statistic | 0.37 | 0.09 | 0.50 | 2.48 | 1.02 | 1.31 |
| p-value | 0.54 | 0.77 | 0.48 | 0.12 | 0.31 | 0.25 |
| endogeneity test | | | | | | |
| F-statistic | 6.69 | 7.18 | 6.42 | 2.56 | 0.03 | 0.01 |
| p-value | 0.01 | 0.01 | 0.01 | 0.11 | 0.85 | 0.91 |

# A Data Appendix

***International Bank Lending Data***

Our data for cross-border bank claims come from the unpublished locational bilateral banking data provided by the Bank of International Settlement (BIS). The sample that we use covers the period 2003 to 2006. The BIS provides two different sets of banking statistics: a locational banking statistics and a consolidated banking statistics. The consolidated banking data provides information on international claims of domestic bank head offices on a *consolidated* basis, i.e. the data consolidate financial positions of the bank head office and their foreign affiliates. This data thus provides a detailed picture of the country exposure risk of a reporting countries banking system. Besides the aggregate volume of international claims, the consolidated banking data also provides disaggregated information based on the sector of the borrower: banks; public sector; and non-bank private sector. In contrast to the consolidated banking statistics, the BIS's locational data is *not* consolidated but based on the residence principle (balance-of-payment principle), i.e. they include information on the *gross* on-balance sheet asset and liability *vis-à-vis* non-resident entities (banks and non-banks). The BIS locational banking statistic thus provides information on the actual amount outstanding of cross-border claims held by the domestic banking system. In the empirical setup we will resort to the locational banking data. The reason is that the theoretical model laid out makes predictions concerning the structure of cross-border lending from country A to country B. The information on international financial claims provided in the locational banking statistics allows discriminating between actual cross-border lending and lending from foreign affiliates. This feature of the data makes the locational statistics more appropriate for a structural estimation of the gravity equation derived above.

***Banking Market Data***

We use the *Financial Structure Database*, collected by Thorsten Beck , Asli Demirguc-Kunt and Ross Eric Levine and explained in Beck et al. (2009), to find data related to the BIS reporting country's banking markets. We measure the size of the banking market in debtor country $j$, which is our proxy for the effective number of banks $n_j$ in the debtor country, using total bank assets. In the original dataset this variable is defined as total bank assets divided by GDP. In order to arrive at the level of total bank assets in the debtor country we multiply this variable with GDP. The remaining debtor-country-specific banking data used in the second stage as reported in Table 3 and Table 4 are directly available in the *Financial Structure Database*. The Financial Structure Database is available from `http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/0,,contentMDK:20696167~pagePK:64214825~piPK:64214943~theSitePK:469382,00.html`

Our main source for country-specific bank lending rate is the *World Development Indicator* (WDI) dataset provided by the World Bank. The lending rates in the WDI data are prime lending rates, i.e. interest rates paid by the banks prime customers. For most Euro area countries the WDI data does not provide information on lending rates after 2003. For these countries we collect bank lending rates from the ECB Statistical Data Warehouse (http://sdw.ecb.europa.eu/). The lending rates for the Euro are countries provided by the WDI prior to 2003 and later on by the ECB are not comparable. Therefore we do not rewrite the WDI lending rates with the ECB lending rate. Instead we choose to begin our estimation with the year 2003. In unreported robustness tests, we let the sample run from 1995 to 2002 and used the WDI lending rates throughout. The results which are available on request are basically the same.

***Distance and Distance-Related Variables***

The variables distance, common language and common border are collected from the dyadic CEPII distance dataset available from the CEPII homepage (http://www.cepii.fr/anglaisgraph/bdd/distances.htm).

Information on membership in the same free trade agreement, common currency, same time zone is retrieved from CEPII Gravity Dataset, created by Keith Head, Thierry Mayer and John Ries to be used in Head et al. (2010), which is available from http://www.cepii.fr/anglaisgraph/bdd/gravity.htm. The variable common legal origin has been constructed from the dataset constructed by Rafael La Porta, Florencio Lopez-de-Silanes, Andrei Shleifer and Robert Vishny to be used in the La Porta et al. (1998), and (for instance) available from http://www.economics.harvard.edu/faculty/shleifer/paper.

# B Theoretical Appendix

The probability of a firm in country $i$ choosing bank $k$ in country $j$ was given by the following integral:

$$\mathbf{P}_{jk} = \int_{-\infty}^{\infty} \frac{1}{\sigma} \exp\left(\frac{x-\mu}{\sigma}\right) \left\{ \exp\left[ -\exp\left(\frac{x-\mu}{\sigma}\right) \right] \right\} \prod_{l=1}^{N} \exp\left[ -\exp\left( \frac{\bar{c}_{ij} - \bar{c}_{il} + \sigma \ln n_l + x - \mu}{\sigma} \right) \right] dx$$

Simplifying the equation by setting $y_{il} = \exp(\bar{c}_{il}/\sigma)$ and $\delta = \exp(\frac{x-\mu}{\sigma})$, and adjusting the integral's range, we arrive at the following, simplified integral:

$$
\begin{aligned}
\mathbf{P}_{jk} &= \int_0^{\infty} \exp(-\delta) \prod_{l=1}^{N} \exp\left[-(y_{ij} y_{il}^{-1} n_l \delta)\right] d\delta \qquad \text{with } k \notin n_j \\
\mathbf{P}_{jk} &= \int_0^{\infty} \exp(-\delta) \exp\left[-\left(\sum_{l=1}^{N} n_l \frac{(y_{il})^{-1}}{(y_{ij})^{-1}} \delta\right)\right] \\
\mathbf{P}_{jk} &= \int_0^{\infty} \exp(-\delta) \exp\left[-\left(\sum_{l=1}^{N} n_l \frac{(y_{il})^{-1}}{(y_{ij})^{-1}} \delta - \frac{(y_{ij})^{-1}}{(y_{ij})^{-1}} \delta\right)\right] \qquad \text{with } k \in n_j. \\
\mathbf{P}_{jk} &= \int_0^{\infty} \exp\left[\delta \frac{1}{(y_{ij})^{-1}} \left(\sum_{l=1}^{N} n_l (y_{il})^{-1}\right)\right] d\delta
\end{aligned}
$$

The expansion by $\left(-\frac{(y_{ij})^{-1}}{(y_{ij})^{-1}}\delta\right)$ stems from adding an element to the sum that until then was excluded: the firm-bank relationship $jk$. The sum now runs over all elements. Integrating yields

$$
\begin{aligned}
\mathbf{P}_{jk} &= -\frac{(y_{ij})^{-1}}{\sum_{l=1}^{N} n_l (y_{il})^{-1}} \left\{ \exp\left[-\frac{\delta}{(y_{ij})^{-1}} \left(\sum_{l=1}^{N} n_l (y_{il})^{-1}\right)\right]\right\}_0^{\infty} \\
&= \frac{(y_{ij})^{-1}}{\sum_{l=1}^{N} n_l (y_{il})^{-1}} \\
&= \frac{\exp\left(-\frac{\bar{c}_{ij}}{\sigma}\right)}{\sum_{l=1}^{N} n_l \exp\left(-\frac{\bar{c}_{il}}{\sigma}\right)}.
\end{aligned}
$$

In showing that the Gumbel distribution is the limiting distribution, $\mu = \sigma \ln n_l$, and $\sigma = \sigma_\varepsilon$ we refer to Mood et al. (1974): 258-260. We start from the logistic distribution which is very similar to the cumulative normal distribution for values that are sufficiently different from zero and one because it is much easier to manipulate because the cumulative normal distribution does not have a closed form (Anderson et al. (1992):16).

# Table 5: Correltion Matrix

This table shows pairwise correlation coefficients of the variable used in the empirical part. $BA_{ij}$ are total loans from country $j$ to country $i$ measured in million US-Dollars

| | $BA_{ij}$ | distance | contiguity | common language | common legal origin | Regional Trade Agreement | Common Currency | Banking Market $Size_j$ | Interest $Rate_j$ | 3 Bank Concen. Ratio | Return on Assets | Cost Income Ratio | Z-Score | Bank Asset to $GDP_j$ | Bank Credit to GDP $_j$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $BA_{ij}$ | 1.00 | | | | | | | | | | | | | | |
| distance | -0.11 | 1.00 | | | | | | | | | | | | | |
| contiguity | 0.10 | -0.40 | 1.00 | | | | | | | | | | | | |
| common language | 0.05 | -0.06 | 0.20 | 1.00 | | | | | | | | | | | |
| common legal origin | 0.02 | -0.01 | 0.15 | 0.39 | 1.00 | | | | | | | | | | |
| Regional Trade Agreement | 0.07 | -0.68 | 0.22 | -0.05 | 0.00 | 1.00 | | | | | | | | | |
| Common Currency | 0.13 | -0.39 | 0.31 | 0.02 | 0.12 | 0.33 | 1.00 | | | | | | | | |
| Banking Market $Size_j$ | 0.11 | 0.03 | -0.01 | -0.06 | -0.13 | -0.03 | 0.01 | 1.00 | | | | | | | |
| Interest $Rate_j$ | -0.04 | 0.08 | 0.13 | -0.01 | 0.13 | -0.08 | -0.05 | -0.17 | 1.00 | | | | | | |
| 3 Bank Concentration Ratio | -0.03 | -0.17 | -0.01 | -0.11 | -0.09 | 0.21 | 0.05 | -0.31 | -0.22 | 1.00 | | | | | |
| Return on Assets | -0.09 | -0.05 | 0.03 | -0.05 | 0.02 | -0.01 | -0.06 | -0.42 | 0.44 | 0.10 | 1.00 | | | | |
| Cost Income Ratio | 0.03 | -0.03 | 0.05 | 0.02 | -0.01 | 0.09 | 0.02 | 0.28 | -0.07 | 0.13 | -0.41 | 1.00 | | | |
| Z-Score | -0.03 | 0.03 | -0.03 | -0.04 | -0.04 | 0.05 | 0.02 | -0.18 | -0.16 | 0.25 | -0.22 | 0.07 | 1.00 | | |
| Bank Asset to $GDP_j$ | 0.05 | -0.12 | -0.06 | -0.17 | -0.19 | 0.13 | 0.02 | 0.21 | -0.41 | 0.58 | 0.00 | -0.02 | 0.09 | 1.00 | |
| Bank Credit to GDP $_j$ | 0.04 | -0.15 | -0.07 | -0.12 | -0.15 | 0.16 | 0.02 | 0.10 | -0.45 | 0.60 | 0.07 | -0.06 | 0.08 | 0.95 | 1.00 |

## Graz Economics Papers

For full list see:

http://www-classic.uni-graz.at/vwlwww/forschung/RePEc/

Address: Department of Economics, University of Graz,
Universitätsstraße 15/F4, A-8010 Graz

---

03–2012 **Bettina Brüggemann, Jörn Kleinert, Esteban Prieto** The Ideal Loan and the Patterns of Cross-Border Bank Lending

02–2012 **Michael Scholz, Jens Perch Nielsen, Stefan Sperlich**: Nonparametric prediction of stock returns guided by prior knowledge

01–2012 **Ronald Wendner**: Ramsey, Pigou, heterogenous agents, and non-atmospheric consumption externalities