



WP/18/46

IMF Working Paper

The Global Banking Network: What is Behind
the Increasing Regionalization Trend?

by Eugenio Cerutti and Haonan Zhou

IMF Working Papers describe research in progress by the authors and are published to elicit comments and to encourage debate. The views expressed in IMF Working Papers are those of the authors and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

I N T E R N A T I O N A L M O N E T A R Y F U N D

IMF Working Paper

Research Department

**The Global Banking Network: What is Behind the
Increasing Regionalization Trend?**

Prepared by Eugenio Cerutti and Haonan Zhou*

Authorized for distribution by Maurice Obstfeld

March 2018

IMF Working Papers describe research in progress by the authors and are published to elicit comments and to encourage debate. The views expressed in IMF Working Papers are those of the authors and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

Abstract

This paper analyses the nature of the increasing regionalization process in global banking. Despite the large decline in aggregate cross-border banking lending volumes, some parts of the global banking network are currently more interlinked regionally than before the Global Financial Crisis. After developing a simple theoretical model capturing banks' internationalization decisions, our estimation shows that this regionalization trend is present even after controlling for traditional gravitational variables (e.g. distance, language, legal system, etc.), especially among lenders in EMs and non-core banking systems, such as Australia, Canada, Hong Kong, and Singapore. Moreover, this regionalization trend was present before the GFC, but it has increased since then, and it seems to be associated with regulatory variables and the opportunities created by the retrenchment of several European lenders.

JEL Classification Numbers: F34, F36, F39, G15.

Keywords: Cross-border Lending, Banking Network, Financial Regionalization.

Authors' E-Mail Address: ECerutti@imf.org, HZhou@imf.org

* Cerutti and Zhou are both at the Research Department of the International Monetary Fund (IMF). We are grateful to Stijn Claessens, Ricardo Correa, Giovanni Dell'Ariccia, Jana Gieck-Bricco, Tumer Kapan, Catherine Koch, Sole Martinez-Peria, Maury Obstfeld, TengTeng Xu for helpful comments, to Paola Ganum for help with the data from the RES Bank Contagion Module, and to participants at the Central Bank of Ireland conference on "Global Financial Linkages", and an IMF seminar for useful advice. Author's E-mail addresses: ECerutti@imf.org; HZhou@imf.org. The opinions expressed herein are solely the responsibility of the authors and should not be interpreted as reflecting those of the IMF, its Executive Board, or IMF management.

| Contents | Page |
|--|------|
| I. Introduction | 4 |
| II. Literature Review and Contributions | 7 |
| III. A model of heterogeneous bank internationalization | 9 |
| A. Setup | 9 |
| B. Bank's Problem | 11 |
| C. Aggregation | 12 |
| D. Model Estimation | 14 |
| IV. Data and Estimation | 16 |
| V. Baseline Results | 22 |
| VI. Distinguishing Regional Effect from Traditional Gravity Factors | 27 |
| VII. What else is the regional dummy capturing? | 30 |
| VIII. Conclusion | 37 |
| References | 40 |
| Appendix | |
| A. Additional Estimation Results | 44 |
| A.1. Alternative Regional Groupings | 44 |
| A.2. Alternative Methods | 49 |
| A.3. Gravity Factor Interactions, Alternative Definition of Flow | 55 |
| List of Tables | |
| 1. Summary Statistics | 19 |
| 2. List of CBS Reporting Countries, Regional Classification and Grouping | 21 |
| 3. Baseline Regional Grouping: First Stage – Direct Cross-Border | 23 |
| 4. Baseline Regression: First Stage – Local Affiliate | 24 |
| 5. Baseline Estimation: Second Stage – Direct Cross-Border | 25 |
| 6. Baseline Estimation: Second Stage – Local Affiliates | 26 |
| 7. Region as an Interaction Term – First Stage | 28 |
| 8. Region as an Interaction Term – Second Stage | 29 |
| 9. Variable Definitions: Interaction Variables | 32 |
| 10. Summary Statistics: Additional Interaction Variables | 33 |
| 11. Interaction with Bank Activity Restrictions | 34 |
| 12. Interaction with Trade Linkages | 35 |
| 13. Peripheral Lenders: Interaction with Global Banks' Retrenchment | 36 |
| A1. WB Regional Grouping: First Stage – Direct Cross-Border | 44 |

| | |
|---|----|
| A2. WB Regional Grouping: First Stage – Local Affiliate | 45 |
| A3. IMF Regional Grouping: First Stage – Direct Cross-Border | 45 |
| A4. IMF Regional Grouping: First Stage – Local Affiliate | 46 |
| A5. WB Regional Grouping: Second Stage – Direct Cross-Border | 46 |
| A6. WB Regional Grouping: Second Stage – Local Affiliates | 47 |
| A7. IMF Regional Grouping: Second Stage – Direct Cross-Border | 47 |
| A8. IMF Regional Grouping: Second Stage – Local Affiliates | 48 |
| A9. Non-linear Least Squares Regression: Second Stage – Direct Cross Border | 50 |
| A10. Non-linear Least Squares Regression: Second Stage – Local Affiliate | 51 |
| A11. Semiparametric Regression: Second Stage | 52 |
| A12. Poisson Regression | 53 |
| A13. OLS Regression | 54 |
| A14. Region as an Interaction Term: First Stage, 2006Q1–2009Q2 Sample | 56 |
| A15. Region as an Interaction Term: Second Stage, 2006Q1–2009Q2 Sample | 57 |
| A16. OLS Regression: Log Difference | 58 |

“It has been said that arguing against globalization is like arguing against the laws of gravity.”
 – Kofi Annan, opening of the 53rd Annual DPI/NGO Conference, August 28, 2000.

I. INTRODUCTION

The structure of the global banking network substantially changed after the global financial crisis (GFC). The total amount of cross-border banking lending within the global banking network shrunk, driven by a sharp reduction in cross-border lending activity by most European banking systems (European Central Bank, 2014; International Monetary Fund, 2015; McCauley and others, 2017). Yet, at the same time, there was an increase in the number of lender-borrower connections within the same region in the periphery of the banking network (Cerutti and Zhou, 2017). This latter phenomenon is often called regionalization, and is associated with the expansion of cross-border lending by foreign banks, especially the banks outside the main global banking systems (US, Euro Area, UK, and Japan), through their affiliates in neighboring countries (Bank for International Settlements, 2014; Claessens and van Horen, 2015; International Monetary Fund, 2015, 2017), as well as their increasing direct cross-border lending to borrowing neighbors (Cerutti and Zhou, 2017).¹

The purpose of this paper is to go one step beyond documenting or presenting anecdotal evidence on banking regionalization as is currently done in the literature, and to analyze the nature of the regionalization phenomenon itself, including the factors that could be associated with this recent apparent regional trend in both direct cross-border and foreign affiliates’ banking lending. Three main characteristics need to be taken into account in this endeavor. First, the drivers that could be embedded in a regional strategy are multidimensional. Foreign banks’ preference for regional cross-border lending could be capturing traditional geographical and cultural factors (e.g., information asymmetries that are proportional to the distance between borrowers and lenders, as well as legal, language, historical factors common across countries), and/or institutional and idiosyncratic factors (level of bank regulation, follow-your-client motives, international retrenchment of some other international banks after GFC, etc.). Second, numerous lender-borrower links across countries are null, reflecting that borrowers are only linked with very few banking systems. This seems similar to the challenges faced in the trade literature. Silva and Tenreyro (2006) and Helpman, Melitz, and Rubinstein (2008)

¹Cross-border lending by international banks can be executed through two different types of lending: (i) Direct cross-border lending (e.g., the headquarters of a German international bank lending directly to a Brazilian corporation), or (ii) foreign affiliates’ local lending (e.g., the lending from a foreign subsidiary or branch of the German international bank operating in Brazil to the Brazilian corporation).

consider the need to take into account both the intensive margin (characterized by the quantity of flows), and the extensive margin (characterized by lending / no-lending decisions), so as to consistently specify and estimate the gravity equation. Last, but not least, recent literature has highlighted that banks differ substantially in their cross-border expansion strategy due to heterogeneity explained by productivity, size, and profitability differences (Buch, Koch, and Koetter, 2011, 2014; Buch and Lipponer, 2007).

Against this backdrop, we start by building a simple theoretical model of banks' cross-country expansion, borrowing concepts from recent advances in gravity trade models. We consider both the intensive and extensive margins to analyze the level at which regionalization occurs, as well as features of banking models explaining the internalization strategy of banks. Together with country-level characteristics as additional barriers, banks' decision to extend new lending contributes to the null cross-border banking flows often observed throughout international banking statistics. As a result, empirical models that fail to deal with zero flows specifically would yield inconsistent estimates on the determinant of international banking flows. This highly stylized model, in the spirit of Helpman, Melitz, and Rubinstein (2008) and Fillat and others (2017), introduces latent bank heterogeneity to correct for the potentially heterogeneous effect of productivity barriers on banks' global strategy. Meanwhile, the self-selection of banks into global operations due to common gravitational factors is captured by a Heckman-style selection term. Using a two-stage estimation procedure, our framework yields a gravity equation that allows us to investigate the regionalization phenomenon at a global level, explicitly connecting the lending / no-lending decision with the quantity-of-lending decision, and to obtain consistent estimates at both stages using only aggregate country-level bilateral data, as in the case of trade. The framework has sufficient flexibility to enable us to study various modes of bank internationalization, distinguishing between direct cross-border lending and foreign affiliate lending, as well as to explore factors that could be associated with banks' regional tendencies.

The empirical estimation of the model shows that the regionalization trend is present even beyond the part captured in traditional gravitational variables (e.g., distance, language, legal system). The regional trend seems more prevalent at the extensive margin, and it is particularly present among peripheral lenders in EMs and non-core banking systems, such as Australia, Canada, Hong Kong, and Singapore. Moreover, this regionalization trend is not necessarily a new phenomenon, especially in the case of direct cross-border lending. It was present before the GFC, but it has increased since then in the case of peripheral lenders, while it decreased among core lenders during and after the GFC. By accounting for possible synergies with gravitational factors, we find that the regionalization trend seems to be associated

with the dampening of some barriers to cross-border and local lending (captured by geographical distance) at the extensive margin, as well as increasing the role of others (same legal origin) at the intensive margin. Moreover, at least two factors not captured in the traditional gravitational variables seem to be associated with the regionalization trend: i) after the crisis, international banks are more likely to expand their operations into countries with stricter regulations within the same region; and ii) some of the increase in the regionalization trends by non-core lenders seems to be associated with taking advantage of the opportunities created by the retrenchment of several European lenders, both in terms of direct cross-border and affiliate local claims. We did not find strong evidence that trade linkages are associated with the regionalization trend.

Understanding the drivers of the regionalization process is an important additional step in evaluating the potential consequences of the recent evolution of the global banking system. There is no consensus in the literature on its overall impact. While a regionalized structure may reduce the possibility of large global banks exporting credit booms to local market, as shown by [Claessens and van Horen \(2016\)](#) in the pre-crisis period, it can also increase the vulnerability of the global banking network to regional crisis ([International Monetary Fund, 2015](#)). Additionally, according to the foreign banking literature, it may not lead to the best diffusion of banking technologies ([Buch, Koch, and Koetter, 2014](#)) or to an efficient allocation of global saving across countries ([International Monetary Fund, 2015](#)). In this context, our finding that part of the regional trend is driven by banks seeking to enter neighbor countries with tighter regulatory restrictions, is likely signaling that post-crisis regulatory frameworks might not be conducive to a very diverse or inclusive environment. Before the GFC, the prevailing trend was a cross-border expansion of activities toward neighbor host countries with looser bank regulations (in line with the broader findings of [Houston, Lin, and Ma \(2012\)](#)), but the crisis has not only toughened the regulation environment after the GFC (see [Barth, Caprio, Jr, and Levine \(2013\)](#)), it has also reversed the direction (at the extensive margin) with regional banks expanding to more regulated neighbor countries.

The rest of this paper is organized as follows. Section 2 presents a literature review, highlighting the contributions of our paper to the different strands of the literature. Section 3 presents the theoretical gravitational model capturing banks' international expansions. Section 4 introduces data and the empirical counterpart of the theoretical model. In section 5, we report the results from the baseline model and confirm the existence and prevalence of regionalization. In section 6, we explore whether the regionalization phenomenon reflects an additional unique dimension that might be time- or lender-contingent, or does it merely amplify or dampen the asymmetry captured by the existing factors. In section 7, we focus on factors

that might explain the unique features of regionalization (e.g., regulation, trade linkages, and global banks' cross-border retrenchment). The conclusions are presented in section 8, and we present a rich set of robustness checks in the Appendix.

II. LITERATURE REVIEW AND CONTRIBUTIONS

Our paper is related to, and contributes to different strands of the literature. Several variants of gravity models for international finance have been proposed (Coeurdacier and Martin, 2008; Martin and Rey, 2004; Okawa and van Wincoop, 2012) that mainly focus on deriving gravity equation from portfolio choices. Portes and Rey (2005) estimate a gravity model of asset trade using relevant variables in the context of finance, including market capitalization, financial market development and insider trading. Claessens and van Horen (2014b) introduces the concept of competitor remoteness, constructed as the size-weighted geographical distance, into a gravity model and shows that banks take into account third-country competition in their location decisions. Our model treats banks as heterogenous firms providing liquidity services, setting up an analogy between cross-border bank lending and trade. In that sense, we are closer to Buch et al.'s (2011, 2014) gravitational analyses of the role of German banks' productivity differentials in their cross-border lending. Our work is also related to Fillat and others (2017), who focus on the self-selection of global banks into three different modes of operations (domestic banks, foreign branches, and foreign subsidiaries), and develop a model of entry into global banking to match the stylized facts of global banks' operations in the United States.² Although we take advantage of several data adjustments and breakdowns, the type of bilateral country level data with global coverage is similar to what Houston, Lin, and Ma (2012) use in their gravity equations studying the impact of cross-country regulation differences on international banking flows. They find that international banks seem to engage in regulatory arbitrage by shifting cross-border lending towards countries with lower regulation. We find that this regulatory arbitrage trend was also present in the pre-GFC period when considering regional expansion, but this was no longer the case during and after the GFC. Similarly, we also find heterogeneity across lending banking systems, as in Bouvatier and Delatte (2015), who use a generalized gravity model to show that international

²Similar to our paper, Buch, Koch, and Koetter (2014) and Fillat and others (2017) use a partial equilibrium framework. See Niepmann (2015, 2016) for a small open economy general equilibrium model approach. Unlike Fillat and others (2017) and some of the empirical estimations of Buch, Koch, and Koetter (2014), we do not distinguish between foreign branches and foreign subsidiaries. This is driven by the fact that in order to gain global coverage, we must sacrifice broad access to bank level data. We mostly rely on aggregate BIS data, where we can only distinguish between direct cross-border lending and foreign affiliates' (both subsidiaries' and branches') lending.

banking integration displays significant differences in and out of the Euro area, with the latter integration progressing and the former currently in a cycle towards the downside. More recently, [Karolyi, Sedunov, and Taboada \(2017\)](#), using BIS Consolidated Banking Statistics and a gravitational model as part of their estimations, find that recipient countries benefit from cross-border bank lending through improved financial stability. [Brei and von Peter \(2017\)](#) use BIS Locational Banking Statistics to show that similar to gravity models in trade, distance plays a significant role in bank gravity estimations, even if the physical demands of bank lending are less stringent than that of trade.

Our work is also connected with the large empirical literature analyzing the evolution and structure of the global banking system. A large part of the literature has devoted its attention to understanding the changes in the global banking network during and after the GFC. For example, [Ivashina and Scharfstein \(2010\)](#) and [Cetorelli and Goldberg \(2012\)](#) highlight the role of funding variables (e.g., dollar funding shortages or exposure to wholesale funding) in banks' cross-border deleveraging. [Giannetti and Laeven \(2012\)](#), among others, report evidence of a "flight to home" or "flight to core markets" phenomenon, because of lower expected returns abroad or increased risk aversion. While [Aiyar and others \(2014\)](#) highlight the role of higher capital requirements, [Rose and Wieladek \(2014\)](#) and [Cerutti and Claessens \(2017\)](#) also stress how government support policies favoring domestic lending during the crisis can also explain the reduction in banks' cross border lending.

Putting a bit more emphasis in the post-GFC period, several papers have started to highlight the presence of regional focus in international banks' expansion across different dimensions. For example, [Bank for International Settlements \(2014\)](#) and [International Monetary Fund \(2015, 2017\)](#) highlight the regional focus of the cross-border expansion of some banks in Asia and Latin America through anecdotal evidence on the operations of foreign subsidiaries. [Claessens and van Horen \(2014a\)](#) and [Claessens and van Horen \(2015\)](#) present more comprehensive evidence of the regional focus in the expansion of foreign subsidiaries outside Europe, as well as documenting the expansion of direct cross-border lending to non-OECD countries. Recently, [Lane and Milesi-Ferretti \(2017\)](#) examine the development of international financial integration by constructing a comprehensive external balance sheet of each country. In particular, they find that the overall slowdown in the role of global banks masks regional differences, such as the overseas and regional expansion of Chinese and Colombian banks. [McCauley and others \(2017\)](#) stress the importance of using consolidated data (which would include direct cross-border lending, as well as subsidiaries and branch lending) instead of nationality-based locational data, and conclude that the shrinkage of the banking network, from a consolidated perspective, is more like a regional phenomenon, limited to European

banks, and reversing an earlier banking glut based on very high leverage in Europe. Based on a systematic approach to measuring and mapping interconnectedness using network theory, [Cerutti and Zhou \(2017\)](#) confirm the overall shrinking of both direct cross-border and local affiliate lending, but also note that global aggregates mask rich dynamics of the network. They highlight that although major global lenders decreased lending, they remained the core of the network, while some parts of the network are currently more interlinked regionally than before the crisis, driven by an increase in lending by non-core, non-European banks. In this paper, we complement these papers' findings and contribute to this large part of the literature by focusing our attention on the drivers of the regionalization process in the global banking network.

In the rest of this paper, we first build a general gravitational model that captures heterogeneous banks' internationalization process, and derive the empirical counterpart of the model. In estimating the model, we focus our attention on the drivers of the regionalization process in the global banking network.

III. A MODEL OF HETEROGENEOUS BANK INTERNATIONALIZATION

In this section, we lay out a simple model in the spirit of [Helpman, Melitz, and Rubinstein \(2008\)](#) and [Fillat and others \(2017\)](#). This model, directly estimable based on aggregate, bilateral data, forms the framework for our empirical exercise. To maintain tractability while preserving flexibility in estimation, the model characterizes banks' internationalization decision at both the intensive and the extensive margin in a static, partial equilibrium setting.

A. Setup

Consider a country j , populated by N_j banks with monopoly power. In autarky, each bank has access to the following markets:

- A domestic loan market: banks extend loans to domestic borrowers at the rate of $r_{L_j}(L_j)$, where L_j denotes the value of claims.
- A risk-free asset M with risk-free rate r_f . This risk-free asset can be interpreted, for example, as banks' reserve at their central bank.

- A deposit market: banks take deposits from depositors at the rate of $r_D(D)$, where D is the value of deposits.³

The bank has the option to go global in two ways: it can engage in direct cross-border lending by extending cross-border claims through its domestic headquarter, or lend to local borrowers through local affiliates (subsidiaries and/or branches). Each mode of operation x involves paying an upfront, country-pair specific fixed cost, denoted by f_{ij}^x , as well as a monitoring cost that depends on the level of lending. For example, suppose a bank in country j has a direct cross-border claim of L_{ij}^{CB} on country i . Its profit of engaging in direct cross-border lending to country i is

$$\Pi_{ij}^{CB} = \tau_{ij}^{CB} r_{L_{ij}}(L_{ij}^{CB}) L_{ij}^{CB} - C_j(a) L_{ij}^{CB} - c_j f_{ij}^{CB} \quad (1)$$

The first component of Equation (1) is the total return of L_{ij}^{CB} lending. $\tau_{ij} < 1$ (for $i \neq j$) represents the lender-borrower specific cost of lending, following the melting iceberg cost in trade literature. Only τ_{ij}^{CB} fraction of the total gross return $r_{L_{ij}}(L_{ij}^{CB}) L_{ij}^{CB}$ can be “shipped” back to the headquarters. The second term, $C_j(a) L_{ij}^{CB}$, represents the bank’s monitoring cost of the claims. a is bank-specific, drawn from a continuous distribution $G(a)$ with support $[a_L, a_H]$. $C_j(\cdot)$ is a country-specific multiplier, increasing in a , so that the cost of one unit of direct cross-border claim reflects both the efficiency of individual banks (represented by the inverse of a), as well as the overall banking environment of lender countries. Finally, a fixed cost $c_j f_{ij}^{CB}$ has to be charged for the direct cross-border lending service, where c_j is a country-specific cost multiplier similar to $C_j(\cdot)$ above. Unlike $C_j(\cdot)$, c_j is not a function of a .

On the other hand, the profit of a bank conducting local business through foreign affiliates is similarly formulated:

$$\Pi_{ij}^S = \tau_{ij}^S r_{L_{ij}}(L_{ij}^S) L_{ij}^S - C_i(a) L_{ij}^S - c_j f_{ij}^S \quad (2)$$

where the superscript S denotes subsidiaries and/or branches. Since now the claims L_{ij}^S is managed by the local affiliates in the recipient country, the monitoring cost is now $C_i(a)$ rather than $C_j(a)$, reflecting the overall banking environment in the borrower country. The form of Equation (2) assumes that parent bank is able to transfer its efficiency entirely to

³We do not stress the funding side of bank’s balance sheet in our model by separating retail deposits from wholesale funding, as our model does not have predictive contents on the interaction between funding and internationalization decision of banks.

its affiliates.⁴ For the barriers to banking, we can typically assume that $\tau_{ij}^S < \tau_{ij}^{CB}$ and $f_{ij}^S > f_{ij}^{CB}$. The former assumption is mentioned in [Buch, Koch, and Koetter \(2014\)](#) as τ reflects information and regulation cost, intuitively lower when banks conduct lending as entities in borrowing countries. To maintain flexibility, we are not imposing this assumption *a priori* in the estimations. The latter assumption on fixed costs is likely to be true as establishing local affiliates (either in the form of branches or subsidiaries) may involve additional, distinct dimensions of setup costs (see [Cerutti, Dell’Ariccia, and Peria \(2007\)](#) for more details). We allow for this possibility in our estimations for local affiliate flows using a different set of first-stage excluded variables than the direct cross-border version.

B. Bank’s Problem

Banks in country j choose the optimal level of domestic and foreign claims to country i (if they enter), as well as deposits and risk-free asset, by maximizing its operating profit subject to the balance-sheet constraint. In the most general case, the bank’s optimization problem can be written as

$$\max_{L_j, L_{ij}^{CB}, L_i^S, D, M} \Pi_j + \Pi_{ij}^{CB} + \Pi_{ij}^S + r_f M - r_D(D)D \quad (3)$$

$$\text{s.t. } E + D \geq L_j + L_{ij}^{CB} + L_{ij}^S + M \quad (4)$$

where E is the equity endowment of banks. Trivially, we assume $\tau_{jj} = 1$ for all j . If banks do not lend direct cross-border (locally) to country i , then Π_{ij}^{CB} (Π_{ij}^S) = 0.⁵

To get closed-form solutions to the optimization problem, we follow [Fillat and others \(2017\)](#) to assume a country-specific, constant-elasticity loan demand function. More specifically, we model the bank as facing a downward sloping loan demand curve in all three modes of operation (domestic, direct cross-border, local affiliate), and an upward sloping demand curve of deposits, possibly due to its monopolistic power and the ability to transfer its efficiency and market power to its affiliates. Take direct cross-border claim as an illustration. For country i , $L_{ij}^{CB}(r_{L_{ij}^{CB}}) = r_{L_{ij}^{CB}}^{-\epsilon_{CB}} A_i^{CB}$, where A_i^{CB} is the total direct cross-border loan market size, and ϵ_{CB} is the demand elasticity of direct cross-border claims. Similarly, deposit demand function

⁴We can also assume that parent bank is only able to transfer a constant fraction of its efficiency, but this is trivial as the fraction eventually enters as the intercept in the estimation equation after log-linearization.

⁵Our modeling of the bank’s problem is a modified version of the Monti-Klein model of a monopolistic bank, extended by [Fillat and others \(2017\)](#). [Freixas and Rochet \(2008\)](#) provides an introduction to the original model. We simplify the extended model by removing the components of deposit insurance and capital requirement, noting that these components do not alter the first-order conditions.

is modeled as $D(r_D) = r_D^\theta B_j$, where B_j is the total deposit market size.⁶ We drop country subscript in the deposit demand function as the function is only applicable to lender country j .⁷

Use the first order conditions and the demand functions, we can solve the optimization problem. In particular, the solutions to direct cross-border and local affiliate flows from one bank in country j to country i are

$$L_{ij}^{CB*} = \left(\frac{\varepsilon^{CB}}{\varepsilon^{CB} - 1}\right)^{-\varepsilon^{CB}} (\tau_{ij}^{CB})^{\varepsilon^{CB}} (r_f + C_j(a))^{-\varepsilon^{CB}} A_i^{CB} \quad (5)$$

$$L_{ij}^{S*} = \left(\frac{\varepsilon^S}{\varepsilon^S - 1}\right)^{-\varepsilon^S} (\tau_{ij}^S)^{\varepsilon^S} (r_f + C_i(a))^{-\varepsilon^S} A_i^S \quad (6)$$

We add one additional parametric assumption to simplify the second-stage derivation. This assumption establishes a linear relationship between variable cost and fixed cost multiplier.

Assumption 1. $C_j(a) = ac_j - r_f, ac_j > r_f \forall j$.

As a direct result of this assumption, we have two equilibrium loan equations that are purely multiplicative.

$$L_i^{CB*} = \left(\frac{\varepsilon^{CB}}{\varepsilon^{CB} - 1}\right)^{-\varepsilon^{CB}} (\tau_{ij}^{CB})^{\varepsilon^{CB}} (ac_j)^{-\varepsilon^{CB}} A_i^{CB} \quad (7)$$

$$L_i^{S*} = \left(\frac{\varepsilon^S}{\varepsilon^S - 1}\right)^{-\varepsilon^S} (\tau_{ij}^S)^{\varepsilon^S} (ac_i)^{-\varepsilon^S} A_i^S \quad (8)$$

C. Aggregation

From the optimization problem, one observation is immediate: Optimal quantity decisions on various types of claims, conditional on lending, are independent of each other, in particular independent of domestic lending decision. Therefore, banks extend direct cross-border (local) claims to country i if and only if the *additional* profit generated by each type of instrument Π_{ij}^{CB} (Π_{ij}^S) is positive. As a result, there exists a cutoff inverse efficiency level a_{ij}^{CB} and a_{ij}^S for each type of foreign claims, at which banks are indifferent between lending / no-lending. For example, banks are willing to extend direct cross-border claims to country i if $a < a_{ij}^{CB}$. Thus,

⁶We abstract from deriving the loan demand function, noting that it is technically straightforward by introducing household and production and augmenting the model.

⁷It is worth noting that the simplified balance sheet constraint is given from the perspective of the headquarter. Thus, local affiliates funded by domestic deposits are excluded from consideration. We take account of this fact as we adjust the aggregate data used in the estimation by deposit and loan ratio.

we are not able to observe direct cross-border operation of country j on country i if $a_{ij}^{CB} < a_L$, due to the fact that even the most efficient banks in country j would fail to secure positive profits by engaging in direct cross-border lending.

To arrive at observable country-level aggregate claims, let $\int_{a_L}^{a_{ij}^x} L_i^*(a) dG(a)$ be the aggregate type- x claims on country i of banks in country j . Following [Helpman, Melitz, and Rubinstein \(2008\)](#), we make the following assumption on the distribution of the latent parameter a :

Assumption 2. $\frac{1}{a}$ follows a Pareto distribution with shape parameter $k, k > \max\{\varepsilon^{CB}, \varepsilon^S\}$. The distribution is truncated to the support $[a_L, a_H]$.

By assumption, the distribution function $G(a) = \frac{a^k - a_L^k}{a_H^k - a_L^k}$.⁸ Evaluate the integral and multiply by the total number of banks, we have the total direct cross-border claims, CB_{ij} , can be expressed as

$$\begin{aligned} CB_{ij} &= N_j \int_{a_L}^{a_{ij}^{CB}} L_i^{CB*}(a) dG(a) \\ &= N_j \tau_{ij}^{\varepsilon^{CB}} \left(\frac{\varepsilon^{CB}}{\varepsilon^{CB} - 1} \right)^{-\varepsilon^{CB}} c_j^{-\varepsilon^{CB}} A_i^{CB} \frac{k a_L^{k-\varepsilon^{CB}}}{(a_H^k - a_L^k)(k - \varepsilon^{CB})} W_{ij}^{CB} \end{aligned} \quad (9)$$

where $W_{ij}^{CB} = \max\left\{\left(\frac{a_{ij}^{CB}}{a_L}\right)^{k-\varepsilon^{CB}} - 1, 0\right\}$.

We now have a purely multiplicative model which can be log-linearized. Following the notation of [Helpman, Melitz, and Rubinstein \(2008\)](#), we assume a multiplicative form of variable costs τ_{ij}^{CB} :

Assumption 3. $\tau_{ij}^{\varepsilon^{CB}} = D_{ij}^{\gamma^{CB}} e^{-u_{ij}^{CB}}$, where D_{ij} represents the symmetric distance between i and j and $u_{ij}^{CB} \sim \mathcal{N}(0, \sigma_{u^{CB}}^2)$, representing an i.i.d. unobserved trade friction. Similar assumption is imposed for $\tau_{ij}^{\varepsilon^S}$.

We thus have the estimation equation for direct cross-border claims (local affiliate claims equation are similarly formulated). Lowercase letters denote logged variables.

$$cb_{ij} = \beta_0 + \lambda_j^{CB} + \chi_i^{CB} + \gamma d_{ij} + w_{ij}^{CB} - u_{ij}^{CB} \quad (10)$$

⁸The assumption of Pareto distribution of inverse bank efficiency seems even more appropriate in banking than in trade. Using a bank-level dataset from Germany, [Niepmann \(2016\)](#) estimates the overhead cost distribution via maximum likelihood. The distribution well approximates a Pareto distribution. In a related paper, [Bremus and others \(2014\)](#) build a general equilibrium model with a heterogenous banking sector, in which bank efficiency is Pareto distributed. They use the model to examine the granularity effect of bank size on macroeconomic outcomes.

with λ_j^{CB} and χ_i^{CB} being respectively lender and borrower fixed effects. In the actual estimation, d_{ij} will be proxied by a set of gravity factors, while w_{ij}^{CB} will be estimated from the first stage. OLS estimation, on the other hand, ignores w_{ij}^{CB} , resulting in inconsistent estimation of the coefficients. A similar formulation can be derived for local affiliate lending.

D. Model Estimation

To estimate w_{ij}^{CB} (w_{ij}^S), we start from the extensive margin and follow the procedure described by [Helpman, Melitz, and Rubinstein \(2008\)](#). It is apparent that since domestic/global decisions are independent, banks extend direct cross-border or local affiliates claims if the *additional profit* from lending is positive. Two latent variables, Z_{ij}^x , where $x \in \{CB, S\}$, can be defined based on the most efficient banks' productivity such that

$$Z_{ij}^x = \frac{\Pi_{ij}^{*x}(a_L)}{c_j f_{ij}^x} = \frac{\left(\frac{\varepsilon^x}{\varepsilon^x - 1}\right)^{-\varepsilon^x} (a_L c_j)^{-\varepsilon^x} \tau_{ij}^{\varepsilon^x} A_i^x \left(\frac{a_L c_j}{\varepsilon^x - 1} + r_f\right)}{c_j f_{ij}^x} \quad (11)$$

Then, the most efficient banks enter into direct cross-border lending (local affiliate lending) if $Z_{ij}^{CB} > 1$ ($Z_{ij}^S > 1$). This is exactly when we would observe positive direct cross-border (local) flows. In the case of direct cross-border lending, let $\frac{ac_j}{\varepsilon^{CB} - 1} + r_f = \exp(l_j^{CB})$. Model fixed costs by $f_{ij}^{CB} \equiv \exp(\phi_{1j}^{CB} + \phi_{2i}^{CB} + \kappa \phi_{ij}^{CB} - v_{ij}^{CB})$, where $v_{ij}^{CB} \sim \mathcal{N}(0, \sigma_{v^{CB}}^2)$, respectively. The ϕ_j variables denote fixed financial barrier imposed by origin country on all recipients; the ϕ_i variables, accordingly, represent fixed financial barrier imposed by recipient on all lenders. The ϕ_{ij} variables are observed measures of any *additional* country-pair specified fixed banking costs. Note that it is possible that the types of fixed cost barriers for setting up cross-border / local subsidiaries lending differ, so we allow for this difference by putting superscripts on all explanatory variables of the fixed cost. With a similar formulation for local affiliate lending, taking log and recall Assumption 3, we have the latent variables z_{ij}^x , where $x \in \{CB, S\}$, can be expressed as

$$z_{ij}^x = \gamma_0^x + \xi_j^x + \zeta_i^x + \gamma^x d_{ij} - \kappa^x \phi_{ij}^x + \eta_{ij}^x \quad (12)$$

where $\eta_{ij}^x \equiv u_{ij}^x + v_{ij}^x \sim \mathcal{N}(0, \sigma_{u^x}^2 + \sigma_{v^x}^2)$ is i.i.d., and apparently correlated with u_{ij} in the second stage equation.⁹ Equation (12) thus gives the foundation for first-stage estimation using probit.

⁹Note that here, similar to [Helpman, Melitz, and Rubinstein \(2008\)](#), we assume that v^x and u^x are homoskedastic. According to [Silva and Tenreiro \(2015\)](#), this distributional assumption is key to establishing the framework.

It is now straightforward to estimate the first stage selection equation. Let $x \in \{CB, S\}$. Define the indicator variable T_{ij}^x to be 1 whenever we observe positive lending of type x from country j to country i and 0 when we do not. Following [Helpman, Melitz, and Rubinstein \(2008\)](#), we specify the following Probit equation:

$$\begin{aligned} \rho_{ij}^x &:= \Pr(T_{ij}^x = 1 | \text{observed variables}) \\ &= \Phi(\gamma_0^{x*} + \xi_j^{x*} + \zeta_i^{x*} + \gamma^{x*} d_{ij} - \kappa^{x*} \phi_{ij}^x) \end{aligned} \quad (13)$$

where the asterisk denotes standardization, i.e. division by the standard error σ_{η^x} .

In the first-stage estimation, two terms are of particular interest:

- Let $\hat{\rho}_{ij}^x$ be the predicted probability that country j establishes a type- x connection to country i . Let $\hat{z}_{ij}^{x*} = \Phi^{-1}(\hat{\rho}_{ij}^x)$ be the predicted value of the latent variable $z_{ij}^{x*} \equiv \frac{z_{ij}^x}{\sigma_{\eta^x}}$. Then W_{ij}^x (recall Equation (9)) can be obtained from

$$W_{ij}^x = \max\{(Z_{ij}^{x*})^\delta - 1, 0\} \quad (14)$$

where $\delta = \sigma_\eta(k - \varepsilon^x)/\varepsilon^x$.¹⁰ The term W_{ij} captures the unobserved bank-level heterogeneity that affects the decision to go global.

- Country pairs' selection into banking partners is also controlled by the Heckman selection term, as this term reflects selection into lending/borrowing due to country-pair characteristics. To derive consistent estimates, we would like to know the expectation of w_{ij}^x , z_{ij}^x and u_{ij}^x conditional on regressors and entry, i.e. $T_{ij}^x = 1$. We start with consistently estimating the conditional expectation of η_{ij}^{x*} . Define

$$\begin{aligned} \bar{\eta}_{ij}^{x*} &:= \mathbb{E}[\eta_{ij}^{x*} | \cdot, T_{ij}^x = 1] = \mathbb{E}[\eta_{ij}^{x*} | \gamma_0^{x*} + \xi_j^{x*} + \zeta_i^{x*} + \gamma^{x*} d_{ij} - \kappa^{x*} \phi_{ij}^x + \hat{\eta}_{ij}^{x*} > 0] \\ &= \mathbb{E}[\eta_{ij}^{x*} | \eta_{ij}^{x*} > -z_{ij}^{x*}] = \frac{\phi(z_{ij}^{x*})}{\Phi(z_{ij}^{x*})} \end{aligned} \quad (15)$$

with the last equality due to the fact that η_{ij}^{x*} is standard normal. Therefore, a consistent estimation of $\bar{\eta}_{ij}^{x*}$ can be defined as $\hat{\eta}_{ij}^{x*} = \frac{\phi(\hat{z}_{ij}^{x*})}{\Phi(\hat{z}_{ij}^{x*})}$, by plugging in the predicted value of z_{ij}^{x*} into the inverse Mills ratio function.

Recall that

$$\begin{bmatrix} u_{ij}^x \\ \eta_{ij}^{x*} \end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \begin{bmatrix} \sigma_{u^x}^2 & \sigma_{u^x} \eta^{x*} \\ \sigma_{u^x} \eta^{x*} & 1 \end{bmatrix}) \quad (16)$$

¹⁰To derive this, note that $Z_{ij}^x = (\frac{a_{ij}^x}{a_{ij}^x})^{-\varepsilon^x}$ by the breakeven condition and Equation (11).

where $\sigma_{u^x \eta^{x^*}}$ denotes the covariance of u^x and η^{x^*} . As a result,

$$\mathbb{E}[u_{ij}^x | \cdot, T_{ij}^x = 1] = \text{corr}(u_{ij}^x, \eta_{ij}^{x^*}) \frac{\sigma_{u^x}}{\sigma_{\eta^x}} \bar{\eta}_{ij}^{x^*} \quad (17)$$

Meanwhile,

$$\mathbb{E}[z_{ij}^{x^*} | \cdot, T_{ij}^x = 1] = z_{ij}^{x^*} + \bar{\eta}_{ij}^{x^*} \quad (18)$$

Therefore, the consistent estimate of $\mathbb{E}[z_{ij}^{x^*} | \cdot, T_{ij}^x = 1]$ would be $\hat{z}_{ij}^{x^*} + \hat{\eta}_{ij}^{x^*}$. The estimate of $\mathbb{E}[w_{ij}^{x^*} | \cdot, T_{ij}^x = 1]$ would therefore be $\ln\{\exp[\delta(\hat{z}_{ij}^{x^*} + \hat{\eta}_{ij}^{x^*})] - 1\}$.¹¹

We summarize the estimation strategy below. For $x \in \{CB, S\}$,

- **First stage:** estimate Probit equation

$$\rho_{ij}^x = \Phi(\gamma_0^{x^*} + \zeta_j^{x^*} + \zeta_i^{x^*} + \gamma^{x^*} d_{ij} - \kappa^{x^*} \phi_{ij}^x) \quad (19)$$

with error term $\eta_{ij}^{x^*}$. Get predicted probabilities $\hat{\rho}_{ij}^x$ and predicted latent variable $\hat{z}_{ij}^{x^*} = \Phi^{-1}(\hat{\rho}_{ij}^x)$. Derive consistent estimate of inverse Mills ratio $\hat{\eta}_{ij}^{x^*} = \frac{\phi(\hat{z}_{ij}^{x^*})}{\Phi(\hat{z}_{ij}^{x^*})}$.

- **Second stage:** estimate the non-linear equation

$$x_{ij} = \beta_0^x + \lambda_j^x + \chi_i^x + \gamma^x d_{ij} + \ln\{\exp[\delta^x(\hat{z}_{ij}^{x^*} + \hat{\eta}_{ij}^{x^*})] - 1\} + \beta_{u^x \eta^x} \hat{\eta}_{ij}^{x^*} + e_{ij}^x \quad (20)$$

where $\beta_{u^x \eta^x} \equiv \text{corr}(u_{ij}^x, \eta_{ij}^{x^*}) \frac{\sigma_{u^x}}{\sigma_{\eta^x}}$, e_{ij}^x is an i.i.d. error term and $\mathbb{E}[e_{ij}^x | \cdot, T_{ij} = 1] = 0$. The parameters to estimate are $\beta_0^x, \gamma^x, \delta^x$ and $\beta_{u^x \eta^x}$, with the main parameter of interest being γ^x .

IV. DATA AND ESTIMATION

We now proceed to estimate the empirical counterpart of the model. For bank flows, we use confidential data from the Bank for International Settlements (BIS) Consolidated Banking

¹¹ [Silva and Tenreyro \(2015\)](#) argue that the estimate of $\mathbb{E}[w_{ij}^{x^*} | \cdot, T_{ij}^x = 1]$ by plugging in sample analogs is not consistent due to non-linearity. However, they also suggest that in practice, [Helpman, Melitz, and Rubinstein \(2008\)](#)'s estimator is close to being consistent. In the Appendix (section A.2), we check the robustness of our baseline results, using a semi-parametric approach to relax our parametric assumptions. We also provide the regression results from the Poisson pseudo-maximum likelihood (PPML) estimation, first suggested by [Silva and Tenreyro \(2006\)](#). This procedure has been adopted by [Karolyi and Taboada \(2015\)](#), for the analysis of regulatory arbitrage and cross-border bank acquisitions in the presence of zero incidences in the data.

Statistics (CBS) at ultimate risk basis, covering 28 BIS reporting countries¹² and over 160 non-reporting countries from 2005Q1 to 2016Q2. The CBS data capture the consolidated claims of internationally active banks headquartered in BIS reporting countries, excluding intragroup positions.¹³ Local affiliate claims are adjusted using the methodology of [Cerutti \(2015\)](#). This procedure not only provides a more accurate picture of bilateral banking linkages by excluding local deposit funding from local affiliate claims, but also adapts to the headquarter-based balance sheet constraint posed in our model. Following [Cerutti and Zhou \(2017\)](#), we define banking flow from reporting country j to its counterparty i at time t as the positive difference between the exposure of country j to i from time $t - 4$ to t . Doing so help adjust for seasonality of banking flows while dampening short-term exchange rate variation in influencing the quantity of flows. Formally, let X denote direct cross-border (local) exposure. Then we have,

$$Flow_{ijt}^X = \max(\Delta_4 X_{ijt}, 0) = \max(X_{ijt} - X_{ij(t-4)}, 0) \quad (21)$$

$$Connect_{ijt}^X = I(Flow_{ijt} > 0) \quad (22)$$

where $I(\cdot)$ denotes the indicator function. For each type of flow in $\{CB, S\}$, $Connect_{ijt}$ is used as the dependent variable in the first stage, while the log of $Flow_{ijt}$ is the dependent variable in the second stage. Zeroes in $Flow_{ijt}$ will thus be missing after log transformation, and we observe the second stage dependent variable if and only if the dependent variable in the first stage is equal to one. The first stage dependent variable, $Connect^X$, captures both the “persistent and flickering zeroes” ([Anderson and Yotov, 2017](#)) in the bilateral banking statistics that include no exposure throughout the sample, and temporary halt of new net loan issuances. In this sense, following our model specification, we would be evaluating, in the first stage, whether or not a lender has provided new net credit within a year. We suspect, as is also frequent in the trade literature, that selection bias would arise when there is no new net credit during a specific period.¹⁴ The second stage dependent variable, $Flow^X$, measures the dollar intensity of lender-borrower linkages.

Table 1 contains a summary of the main variables used in our sample. We use a set of standard gravity factors provided by CEPII ([Head and Mayer, 2013](#); [Head, Mayer, and Ries,](#)

¹²South Korea started reporting CBS data in 2013. It is used as a non-reporting country in our sample, with its lending removed to maintain a constant lender composition.

¹³BIS also provides the Location Banking Statistics (LBS). Resident-based, the LBS dataset may double-count the exposure of key financial intermediary countries, and is thus less able to track the ultimate sources and destinations of financial flow. See [Cerutti and Zhou \(2017\)](#) for more details on the comparison between CBS and LBS, and [McCauley and others \(2017\)](#) for further evidence that using consolidated BIS data is preferable when analyzing the evolution of the global banking network.

¹⁴As shown in Section A.3.2 of the Appendix, using the growth rate of exposure (expressed in stock) would not change our findings with respect to regionalization, but will not fit well the gravitational framework.

2010), including common official language, common legal system, as well as indicators for former colonial relationship, common colonizers and log distance. A free trade agreement dummy is also included to detect any synchronization of banking flows and trade finance.¹⁵ Lender (borrower) fixed effects are included in all regressions to account for any time-invariant, unobserved lender (borrower) specific characteristics. Estimating a sample with time dimension, we also include quarter fixed effects to control for the effect of time-varying global factors, despite the fact that our model is static.¹⁶ At both stages, we cluster the standard errors at the conservative country-pair level, similar to [Houston, Lin, and Ma \(2012\)](#).

[Helpman, Melitz, and Rubinstein \(2008\)](#) originally propose using non-linear least squares to estimate the second stage equation ((20)). This method, as observed by [Gómez-Herrera \(2013\)](#), however, is difficult to estimate. This issue is exacerbated as our dataset contains multiple periods and a large set of dummy variables. To increase the efficiency, we exploit the distributional assumption of error term e_{ij}^x in ((20)) and use maximum likelihood to estimate the second stage. In particular, we assume that $e_{ij}^x \sim \mathcal{N}(0, \sigma^{x2})$ and find the value of the parameter γ^x , δ^x and $\beta_{i^x \eta^x}$ and σ^x that maximize the joint likelihood. Compared to the non-linear least squares procedure, we are able to significantly reduce the time toward convergence and obtain very similar results. As we show in Appendix A.2, the point estimates obtained from both procedures are the same.

We include a dummy variable, common region, as our primary variable of interest. The value of this binary variable is set according to geographical definition of continental regions. We use the United Nations geoscheme (M49 standard) in our baseline regression, and confirm the robustness of our results using the regional classification of the World Bank and IMF in the Appendix. To detect if the dynamics of regionalization and bank gravity experience any change due to the crisis, we also report results obtained by interacting the common region dummy with crisis and post-crisis indicator, with the former spanning 2008Q1 to 2009Q2, and the latter 2009Q3–2016Q2, following NBER definition.

Finally, we could further break down the effect of common region by lender groups, as the unbalanced structure of global banking may result in distinct segments of global and regional lenders. Table 2 summarizes the classification of lenders into groups and the regional affiliation of each reporting country under three different schemes. Following [Cerutti and Zhou](#)

¹⁵Our specification is robust to adding a Euro Area dummy that accounts for the effect of a common currency and the dense financial linkages within the Euro Area.

¹⁶Time-invariant gravity factors can be identified from the model as the model does not include country-pair fixed effects. The inclusion of quarter fixed effects, however, preclude the intercept β_0 from being identified. We drop β_0 in all of our estimations.

Table 1. Summary Statistics

| VARIABLES | (1) N | (2) Mean | (3) SD | (4) Min | (5) Max | (6) Source | (7) Country/Pair |
|------------------------------------|----------|-------------|-----------|-----------------------|------------|---------------|---------------------|
| Dependent Variables | | | | | | | |
| Direct Cross-Border Flow | 201492 | 0.336 | 4.000 | 0 | 702.3 | | |
| Direct Cross-Border Connection | 201492 | 0.237 | 0.425 | 0 | 1 | BIS | 5597 |
| Local Affiliate Flow | 201492 | 0.097 | 2.553 | 0 | 297.6 | | |
| Local Affiliate Connection | 201492 | 0.0670 | 0.250 | 0 | 1 | | |
| Instruments | | | | | | | |
| Lender Concentration (%) | 246 | 67.32 | 19.63 | 27.96 | 100 | | 29 |
| Borrower Concentration (%) | 1306 | 71.21 | 19.79 | 22.08 | 100 | World Bank | 163 |
| Lender Overhead % Total Assets | 250 | 1.637 | 1.134 | 0.071 | 9.279 | | 29 |
| Borrower Overhead % Total Assets | 1504 | 3.634 | 3.531 | 3.64×10^{-5} | 83.314 | | 182 |
| Lender Entry Cost | 252 | 7.452 | 11.46 | 0 | 78.4 | | 28 |
| Borrower Entry Cost | 1584 | 45.48 | 92.86 | 0 | 1314.6 | CEPII | 179 |
| Lender Entry Time | 252 | 23.58 | 35.78 | 5.5 | 168 | | 28 |
| Borrower Entry Time | 1584 | 43.01 | 56.73 | 1.5 | 703.5 | | 179 |
| Gravity Factors | | | | | | | |
| Common Region (UN) | 5597 | 0.217 | 0.412 | 0 | 1 | | 5597 |
| Official Language | 5539 | 0.141 | 0.348 | 0 | 1 | | 5539 |
| Common Colonizer (Post-1945) | 5539 | 0.0301 | 0.171 | 0 | 1 | | 5539 |
| Legal System | 5539 | 0.339 | 0.474 | 0 | 1 | | 5539 |
| Colonial Relationship (Historical) | 5423 | 0.036 | 0.187 | 0 | 1 | CEPII | 5423 |
| Free Trade Agreement (WTO) | 48807 | 0.210 | 0.407 | 0 | 1 | | 5423 |
| Log Geographical Distance | 5597 | 8.700 | 0.816 | 4.175 | 9.900 | | 5597 |
| Log Lender GDP ($t - 4$) | 261 | 6.49 | 1.33 | 2.74 | 9.80 | | 29 |
| Log Borrower GDP ($t - 4$) | 1655 | 3.31 | 2.43 | -3.82 | 9.73 | IMF | 185 |

Notes: Definitions of flow and connection follows Equation (21) and Equation (22). Sample period is restricted to 2006-2014. N refers to the number of unique observations, in its original frequency (quarterly for dependent variables, annual for instruments, GDP and FTA, and time-invariant for other gravity factors) at country level or country-pair level for which data is not missing. Here, “lender” refers to a BIS reporting country, while “borrower” refers to a country which is either BIS-reporting or non-reporting. The last column, “Country/Pair”, refers to the number of countries / country pairs for which data is not missing.

(2017), we classify France, Germany, Japan, UK and US as five “core” global lenders, as they are the perennial top-five countries among all reporting countries in terms of total claims to the world. Non-core BIS reporting countries outside Europe are classified as “peripheral” lenders. This group include both advanced economies (Australia, Canada) not traditionally known as global lenders, as well as a number of emerging market economies (Chile, Mexico, among others).¹⁷

¹⁷For robustness, we also include an Euro Area dummy to account for the dense financial connections between countries in the currency union. The results are very similar to our baseline estimates.

Identification of the second stage requires the inclusion of fixed-cost barrier ϕ_{ij}^x in the first stage, which represents the additional country-pair specific fixed cost shifter. The inclusion of Heckman selection term and bank heterogeneity term in the second stage raises a potential multicollinearity problem that may render the model underidentified. As a result, the additional fixed-cost shifter needs to be carefully selected, so that it is not correlated with the variable cost of holding direct cross-border or local claims, and meanwhile facilitates the identification of the second-stage equation. ϕ_{ij}^x is excluded from the second-stage estimation. Thus it is natural to denote this variable as the “instrumental variable”. We explore a wide set of IV candidates, taking into account the potentially different fixed cost barriers faced by banks to engage in direct cross-border lending or local subsidiary operation. Common choices for both types of flows include bank concentration and overhead cost measure from [Beck, Demirguc-Kunt, and Levine \(2000, 2009\)](#); [Cihak and others \(2012\)](#). Both measures are aggregated from Bankscope and updated by the World Bank on an annual basis. Bank concentration captures the share of total assets for the top-three banks in each country, whereas overhead costs include cost items on human resource, office space, IT, etc.. Since setting up local affiliates amounts to setting up overseas firms, we also include measures of the cost, times and procedures for firm entrance from [Head and Mayer \(2013\)](#); [Head, Mayer, and Ries \(2010\)](#) for regressions with local affiliate lending as independent variable. Following [Helpman, Melitz, and Rubinstein \(2008\)](#), for overhead cost, entry cost and entry time, we construct indicators that equal one if both the lender and the borrower country are above the median level, and zero otherwise, to arrive at our final excluded variables. We construct the bank concentration IV by multiplying the concentration of the lender with that of the borrower, and scale to percentages.

It is worth further discussing the choice of excluded variables and the empirical counterpart of our estimation framework. The variables we constructed using the candidates intuitively capture the complementarity of fixed cost barriers between lenders and borrowers, and thus isolate the country-pair level impact from the home-level or host-level impact embodied in country fixed effects that might directly affect variable cost and quantity of flows. Indeed, all indicators have negligible correlations with direct cross-border and local affiliate flows and exposures, intuitively due to the fact that financial transactions, unlike trade flows, demand little physical space other than established infrastructures that facilitate transmission. The measure of overhead costs to total assets typically includes only the items directly related to the fixed expense of bank operations, such as wages, office rentals and IT costs, so as to serve as natural fixed cost shifters. In addition, according to [Niepmann \(2016\)](#), since overhead costs are independent of funding costs, they serve as the ideal proxy for banks’ monitoring technology and thus productivity in models like ours. Borrower bank concentration

Table 2. List of CBS Reporting Countries, Regional Classification and Grouping

| Country | Region (UN) | Region (WB) | Region (IMF) |
|------------------------------|--------------------|---------------------------|-------------------------|
| <u>Core</u> | | | |
| France | Europe | Europe | Central Asia & European |
| Germany | Europe | Europe | Central Asia & European |
| Japan | Asia | East Asia & Pacific | Asia-Pacific |
| United Kingdom | Europe | Europe | Central Asia & European |
| United States | Americas | North America | Western Hemisphere |
| <u>Other European</u> | | | |
| Austria | Europe | Europe | Central Asia & European |
| Belgium | Europe | Europe | Central Asia & European |
| Denmark | Europe | Europe | Central Asia & European |
| Finland | Europe | Europe | Central Asia & European |
| Greece | Europe | Europe | Central Asia & European |
| Ireland | Europe | Europe | Central Asia & European |
| Italy | Europe | Europe | Central Asia & European |
| Luxembourg | Europe | Europe | Central Asia & European |
| Netherlands | Europe | Europe | Central Asia & European |
| Portugal | Europe | Europe | Central Asia & European |
| Spain | Europe | Europe | Central Asia & European |
| Sweden | Europe | Europe | Central Asia & European |
| Switzerland | Europe | Europe | Central Asia & European |
| <u>Peripheral</u> | | | |
| Australia | Oceania | East Asia & Pacific | Asia-Pacific |
| Brazil | Americas | Latin America & Caribbean | Western Hemisphere |
| Canada | Americas | North America | Western Hemisphere |
| Chile | Americas | Latin America & Caribbean | Western Hemisphere |
| Mexico | Americas | Latin America & Caribbean | Western Hemisphere |
| Panama | Americas | Latin America & Caribbean | Western Hemisphere |
| Hong Kong SAR | Asia | East Asia & Pacific | Asia-Pacific |
| India | Asia | South Asia | Asia-Pacific |
| Singapore | Asia | East Asia & Pacific | Asia-Pacific |
| Taiwan Province of China | Asia | East Asia & Pacific | Asia-Pacific |
| Turkey | Asia | Europe | Central Asia & European |

Notes: Countries with top-five total claims in the BIS CBS data are grouped as core global lenders. All regional classifications use the top level (continental level) of each source.

reflects the difficulty for foreign banks to acquire market share in the borrower country, while lender bank concentration might capture the incentive of home banks to internationalize, in order to expand operation beyond the home market and overperform peers. As considered by [Claessens and van Horen \(2012\)](#), both the lender and borrower market development have an impact on foreign banks' profits from going global. By multiplying home country concentration and host country concentration, this measure puts more emphasis on the country-pair specific effect of market characteristics on foreign banks' initiation of new loans.

V. BASELINE RESULTS

In this section, we discuss the results obtained from our baseline estimation. Table 3 to Table 6 report the regression output at both stages for direct cross-border and local affiliate flows. Traditional gravitational factors explain the increase in banking exposure and the magnitude of flows well, and the estimations yield realistic coefficients. Table 3 and Table 4 suggest that the establishment of direct cross-border and local affiliate connections has strong and positive correlations with the size of borrower countries (proxied by GDP, potentially reflecting borrowers' demand for new loans), trade relationship (proxied by FTA), common legal system and colonial ties. The positive effect of borrower income and colonial roots persists when we move to the second stage, as is revealed by Table 5 and 6. Common language, however, is not a determinant of local affiliate flows at both stages, as its coefficient is small and statistically insignificant. Intuitively, acquiring local affiliates can be regarded as one measure to overcome language and cultural barriers for global banks to engage in deeper-level business operations overseas. Instead, being a colonizer seems to provide a positive boost to the amount of direct cross-border lending at the intensive margin. Geographical distance is, without doubt, an important determinant of international banking at both extensive margin and intensive margin. Both connections and size of flows are positively related to geographical proximity, even though financial flows across distant countries are less bound by physical transport costs as in the case of trade. The elasticity of direct cross-border flows with respect to geographical distance is close to -1, while flows via local affiliate are more sensitive to distance barriers, with an elasticity around -1.4. We show in the Appendix (section A.2) that our gravity model is a good fit for international banking flows. The second-stage non-linear least squares estimates yield an R^2 of around 0.65 for direct cross-border flows, and 0.55 for local affiliate flows.

Excluded variables for both direct cross-border and local affiliate flows turn out to be instruments that perform well in assisting identification, and the estimated coefficients at the first stage are meaningful. For local affiliate connections, the instrument related to the cost of establishing entities in borrowing countries is the most significant. The positive sign of the coefficient is intuitive, implying the existence of substantial barriers to initiating local banking business. Overhead cost and the product of home and host bank concentration, on the other hand, are the more significant instruments for direct cross-border flows, and coefficients for both variables are negative. A high overhead indicator suggests that both the lender and borrower banking systems are relatively inefficient, so as to further hinder operations between both countries. A lower concentration in the lender country may indicate a higher degree of competition that compels banks to discover profit opportunities in the cross-border lending

Table 3. Baseline Regional Grouping: First Stage – Direct Cross-Border

| | (1) No Interactions | (2) Crisis | (3) Peripheral | (4) Peripheral + Crisis |
|-----------------------------------|-------------------------|-------------------------|-------------------------|----------------------------|
| Region | -0.00233 (0.00991) | 0.0568*** (0.0126) | -0.0667*** (0.0119) | 0.0180 (0.0156) |
| Region * Crisis | | -0.0559*** (0.00931) | | -0.0752*** (0.0102) |
| Region * Post-crisis | | -0.0752*** (0.00853) | | -0.112*** (0.00928) |
| Region * Peripheral | | | 0.172*** (0.0206) | 0.0605*** (0.0227) |
| Region * Peripheral * Crisis | | | | 0.0863*** (0.0223) |
| Region * Peripheral * Post-crisis | | | | 0.153*** (0.0212) |
| Language | 0.0243** (0.0114) | 0.0238** (0.0114) | 0.0282** (0.0111) | 0.0282** (0.0111) |
| Colonizer | 0.0362* (0.0191) | 0.0361* (0.0190) | 0.0280 (0.0197) | 0.0277 (0.0196) |
| Legal | 0.0218*** (0.00659) | 0.0220*** (0.00658) | 0.0218*** (0.00658) | 0.0217*** (0.00656) |
| Colonial Relation | 0.0911*** (0.0180) | 0.0902*** (0.0180) | 0.0833*** (0.0175) | 0.0815*** (0.0174) |
| FTA | 0.00663 (0.00974) | 0.00723 (0.00972) | 0.0267*** (0.0101) | 0.0255** (0.0101) |
| Log Distance | -0.0661*** (0.00710) | -0.0662*** (0.00709) | -0.0594*** (0.00697) | -0.0603*** (0.00694) |
| Lender Log GDP | 0.209*** (0.0155) | 0.207*** (0.0155) | 0.209*** (0.0156) | 0.169*** (0.0162) |
| Borrower Log GDP | 0.122*** (0.0119) | 0.108*** (0.0120) | 0.123*** (0.0119) | 0.0899*** (0.0121) |
| Concentration | -0.0498** (0.0232) | -0.0472** (0.0232) | -0.0505** (0.0233) | -0.0452* (0.0232) |
| Overhead | -0.0529*** (0.0109) | -0.0524*** (0.0109) | -0.0561*** (0.0109) | -0.0580*** (0.0108) |
| N | 139216 | 139216 | 139216 | 139216 |
| pseudo R-sq | 0.224 | 0.226 | 0.227 | 0.230 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Notes: This table reports the first-stage Probit regression result on the baseline model, with direct cross-border connection as dependent variable. All coefficients are average marginal effects. Standard errors are clustered at country-pair level.

business. A high concentration in the borrower country, on the other hand, may reflect monopoly power that serves as the impediment to financial openness, thus exerting a downward pressure on direct cross-border connections. The validity of our empirical framework is established by the strong significance of the structural parameters δ and β , which contributes

Table 4. Baseline Regression: First Stage – Local Affiliate

| | (1) No Interactions | (2) Crisis | (3) Peripheral | (4) Peripheral + Crisis |
|-----------------------------------|-------------------------|-------------------------|-------------------------|----------------------------|
| Region | 0.0126 (0.0106) | 0.0201 (0.0125) | -0.0193 (0.0127) | -0.00643 (0.0147) |
| Region * Crisis | | 0.000527 (0.00768) | | 0.00136 (0.00805) |
| Region * Post-crisis | | -0.0121 (0.00847) | | -0.0221** (0.00890) |
| Region * Peripheral | | | 0.104*** (0.0320) | 0.0648* (0.0338) |
| Region * Peripheral * Crisis | | | | -0.00457 (0.0172) |
| Region * Peripheral * Post-crisis | | | | 0.0542** (0.0226) |
| Language | 0.00773 (0.0123) | 0.00766 (0.0123) | 0.0112 (0.0124) | 0.0111 (0.0124) |
| Colonizer | 0.0309 (0.0344) | 0.0310 (0.0344) | 0.00933 (0.0339) | 0.00958 (0.0339) |
| Legal | 0.0226*** (0.00714) | 0.0226*** (0.00714) | 0.0227*** (0.00715) | 0.0227*** (0.00714) |
| Colonial Relation | 0.0689*** (0.0187) | 0.0688*** (0.0186) | 0.0644*** (0.0184) | 0.0640*** (0.0183) |
| FTA | 0.0231** (0.0111) | 0.0230** (0.0110) | 0.0311*** (0.0114) | 0.0306*** (0.0114) |
| Log Distance | -0.0356*** (0.00792) | -0.0357*** (0.00792) | -0.0360*** (0.00785) | -0.0362*** (0.00785) |
| Lender Log GDP | 0.0949*** (0.0166) | 0.0920*** (0.0166) | 0.0950*** (0.0167) | 0.0741*** (0.0175) |
| Borrower Log GDP | 0.0645*** (0.0135) | 0.0594*** (0.0137) | 0.0647*** (0.0135) | 0.0502*** (0.0138) |
| Concentration | 0.0253 (0.0227) | 0.0259 (0.0228) | 0.0281 (0.0228) | 0.0301 (0.0229) |
| Overhead | -0.0194 (0.0137) | -0.0185 (0.0136) | -0.0247* (0.0137) | -0.0274** (0.0136) |
| Entry Cost | -0.0341* (0.0182) | -0.0343* (0.0182) | -0.0354* (0.0194) | -0.0357* (0.0193) |
| Entry Time | -0.00654 (0.00997) | -0.00642 (0.00998) | -0.00775 (0.0102) | -0.00678 (0.0102) |
| N | 91996 | 91996 | 91996 | 91996 |
| pseudo R-sq | 0.261 | 0.261 | 0.263 | 0.264 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Notes: This table reports the first-stage Probit regression result on the baseline model, with local affiliate connection as dependent variable. All coefficients are marginal effects. Standard errors are clustered at country-pair level.

Table 5. Baseline Estimation: Second Stage – Direct Cross-Border

| | (1) No Interactions | (2) Crisis | (3) Peripheral | (4) Peripheral + Crisis |
|-----------------------------------|------------------------|-----------------------|-----------------------|----------------------------|
| Region | -0.0591 (0.0844) | -0.0479 (0.0884) | -0.405*** (0.112) | -0.318*** (0.106) |
| Region * Crisis | | 0.0102 (0.0652) | | -0.0138 (0.0758) |
| Region * Post-crisis | | -0.0296 (0.0743) | | -0.153 (0.102) |
| Region * Peripheral | | | 0.807*** (0.166) | 0.606*** (0.156) |
| Region * Peripheral * Crisis | | | | 0.102 (0.124) |
| Region * Peripheral * Post-Crisis | | | | 0.315** (0.153) |
| Language | 0.157 (0.0966) | 0.152 (0.0965) | 0.190** (0.0959) | 0.188** (0.0958) |
| Colonizer | 0.464*** (0.147) | 0.453*** (0.147) | 0.392*** (0.150) | 0.387*** (0.150) |
| Legal | 0.153*** (0.0593) | 0.150** (0.0593) | 0.150** (0.0587) | 0.148** (0.0588) |
| Colonial Relation | 0.840*** (0.132) | 0.824*** (0.132) | 0.814*** (0.127) | 0.802*** (0.128) |
| FTA | 0.131* (0.0792) | 0.130* (0.0792) | 0.215*** (0.0808) | 0.210*** (0.0807) |
| Log Distance | -0.984*** (0.0788) | -0.966*** (0.0787) | -0.964*** (0.0749) | -0.940*** (0.0758) |
| Lender Log GDP | 0.809*** (0.0921) | 0.809*** (0.0909) | 0.790*** (0.101) | 0.758*** (0.0889) |
| Borrower Log GDP | 0.692*** (0.0796) | 0.674*** (0.0782) | 0.683*** (0.0839) | 0.653*** (0.0780) |
| δ | 0.148*** (0.0572) | 0.157*** (0.0339) | 0.149 (0.107) | 0.211*** (0.0419) |
| β | 0.803*** (0.210) | 0.716*** (0.213) | 0.819*** (0.205) | 0.763*** (0.232) |
| σ | 1.535*** (0.0121) | 1.536*** (0.0121) | 1.531*** (0.0121) | 1.531*** (0.0121) |
| N | 37625 | 37625 | 37625 | 37625 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Notes: This table reports the second-stage maximum likelihood estimation result on the baseline model, with log direct cross-border flow as dependent variable. Standard errors are clustered at country-pair level. δ and β refer to, respectively, the parameter of the non-linear adjustment term and the estimated inverse Mill's ratio term in Equation (20). σ is the estimated standard deviation of the disturbance term, e_{ij} , in Equation (20).

to the explanation of the source of bias in traditional bank gravity estimates. For both direct cross-border flow and local affiliate flow, δ is highly significant, implying that unobserved

Table 6. Baseline Estimation: Second Stage – Local Affiliates

| | (1) No Interactions | (2) Crisis | (3) Peripheral | (4) Peripheral + Crisis |
|-----------------------------------|------------------------|----------------------|----------------------|----------------------------|
| Region | -0.0918 (0.288) | -0.269 (0.315) | -0.318 (0.412) | -0.414 (0.401) |
| Region * Crisis | | 0.235 (0.151) | | 0.208 (0.155) |
| Region * Post-crisis | | 0.215 (0.203) | | -0.0713 (0.260) |
| Region * Peripheral | | | 0.661 (0.686) | 0.0585 (0.722) |
| Region * Peripheral * Crisis | | | | 0.471 (0.397) |
| Region * Peripheral * Post-Crisis | | | | 1.367** (0.572) |
| Language | 0.255 (0.265) | 0.254 (0.266) | 0.283 (0.268) | 0.341 (0.271) |
| Colonizer | 0.673 (0.599) | 0.658 (0.601) | 0.384 (0.638) | 0.379 (0.627) |
| Legal | 0.360* (0.207) | 0.355* (0.212) | 0.387* (0.207) | 0.477** (0.226) |
| Colonial Relation | 1.516*** (0.405) | 1.501*** (0.418) | 1.513*** (0.388) | 1.739*** (0.446) |
| FTA | 0.0889 (0.254) | 0.0836 (0.256) | 0.155 (0.274) | 0.291 (0.292) |
| Log Distance | -1.363*** (0.287) | -1.352*** (0.298) | -1.397*** (0.288) | -1.539*** (0.324) |
| Lender Log GDP | -0.271 (0.305) | -0.284 (0.313) | -0.227 (0.305) | -0.178 (0.318) |
| Borrower Log GDP | 1.322*** (0.330) | 1.327*** (0.324) | 1.304*** (0.328) | 1.419*** (0.329) |
| δ | 5.253*** (1.481) | 5.335*** (1.523) | 5.212*** (1.480) | 4.329*** (1.679) |
| β | 1.387* (0.746) | 1.350* (0.791) | 1.454** (0.730) | 1.982** (0.920) |
| σ | 2.310*** (0.0393) | 2.310*** (0.0393) | 2.308*** (0.0393) | 2.305*** (0.0392) |
| N | 11044 | 11044 | 11044 | 11044 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Notes: This table reports the second-stage maximum likelihood estimation result on the baseline model, with log cross-border flow as dependent variable. Standard errors are clustered at country-pair level. δ and β refer to, respectively, the parameter of the non-linear adjustment term and the estimated inverse Mill's ratio term in Equation (20). σ is the estimated standard deviation of the disturbance term, e_{ij} , in Equation (20).

bank-specific heterogeneity plays an important role in determining observed level of bilateral banking flow. Both types of flows are further influenced by the unobserved country-pair shocks to the fixed cost, as indicated by the significance of β .

Finally, we zero in on the coefficients of common region indicator and interactions. Column (2) of Table 3 and also, to a lower extent, column (2) of Table 4 suggests that the financial crisis imposed a negative shock on within-region linkages, but further decomposition (column (3) and (4)) across lender groups finds that non-core, non-European reporting countries (peripheral lenders) seem to thrive toward their respective regions after the GFC. These countries have played the role of regional lenders before the crisis, as is evident from the positive coefficients of common region and peripheral interactions in column (3) of Table 3 and 4. Their role, moreover, seems to be further consolidated after the crisis. The triple interaction terms in column (4) of both tables are positive and significant, and the sum of three marginal effects (region * post-crisis, region * peripheral, region * peripheral * post-crisis) is also positive. Second-stage results are not as significant than first stage, for the reason that a large proportion of regional effects have already been absorbed in the extensive margin. Still, columns (3) and (4) of Table 5 are in line with the first stage, that the regionalization process for direct cross-border flow accelerated after the crisis, with the peripheral lenders acting as key regional players. Similarly, column (4) of Table 6 provides additional evidence that regionalization of local affiliate flows accelerates after the crisis, driven by peripheral, non-global lenders. Our results, as a whole, thus establish regional preference as a key additional determinant of banking connections for the group of non-core, non-European lenders after the financial crisis. In the Appendix (section A.1), we show that this result is largely robust to alternative region classifications.

VI. DISTINGUISHING REGIONAL EFFECT FROM TRADITIONAL GRAVITY FACTORS

The concept of “common region” represents a multi-layer bilateral linkage not limited to geography. In particular, there exists sufficient overlap between regional proximity at the continental level and other gravity factors used in our baseline regressions. For instance, a considerable number of countries in Africa had historical colonial ties with Belgium and France. While the common region indicator enters the baseline regressions with significance, especially in the extensive margin, the regressions are less capable in answering if the regional effect is distinct on its own, or if it is the synergy between regional proximity and traditional gravity factors that contributes to the significance of the regional dummy and its interactions with temporal indicators and lender groups. In other words, if traditional gravity factors serve as proxies for information asymmetry, as the gravity equations often embody, does the commonality in regional affiliation reflect an additional unique dimension that might be time- or lender-contingent, or does it merely amplify or dampen the asymmetry captured by the existing factors?

Table 7. Region as an Interaction Term – First Stage

| | Full Sample | | | | Post-Crisis Sample | | | |
|----------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| | Direct Cross-Border | | Local Affiliate | | Direct Cross-Border | | Local Affiliate | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Region | -0.317*** (0.00403) | -0.313*** (0.00669) | -0.152** (0.0768) | -0.139* (0.0792) | -0.317*** (0.00350) | -0.314*** (0.00594) | -0.202*** (0.0702) | -0.183** (0.0761) |
| Region * Peripheral | | 0.135*** (0.0238) | | 0.117*** (0.0335) | | 0.150*** (0.0259) | | 0.134*** (0.0366) |
| Language | 0.0322** (0.0128) | 0.0367*** (0.0129) | 0.0265* (0.0151) | 0.0325** (0.0155) | 0.0372** (0.0150) | 0.0421*** (0.0151) | 0.0317* (0.0164) | 0.0377** (0.0169) |
| Colonizer | 0.0536** (0.0241) | 0.0761*** (0.0251) | 0.0444 (0.0677) | 0.0751 (0.0756) | 0.0741*** (0.0281) | 0.100*** (0.0293) | 0.0536 (0.0772) | 0.0901 (0.0878) |
| Legal | 0.0181** (0.00747) | 0.0197*** (0.00750) | 0.0257*** (0.00861) | 0.0278*** (0.00868) | 0.0135 (0.00848) | 0.0153* (0.00851) | 0.0278*** (0.00961) | 0.0305*** (0.00967) |
| Colonial Relation | 0.0943*** (0.0189) | 0.0891*** (0.0188) | 0.0675*** (0.0198) | 0.0615*** (0.0196) | 0.0954*** (0.0220) | 0.0892*** (0.0219) | 0.0761*** (0.0227) | 0.0692*** (0.0225) |
| FTA | 0.0234* (0.0120) | 0.0287** (0.0122) | 0.0277** (0.0137) | 0.0345** (0.0142) | 0.0121 (0.0134) | 0.0180 (0.0136) | 0.0309** (0.0145) | 0.0385*** (0.0149) |
| Log Distance | -0.127*** (0.00921) | -0.113*** (0.00939) | -0.0632*** (0.0105) | -0.0604*** (0.0105) | -0.130*** (0.0104) | -0.114*** (0.0107) | -0.0655*** (0.0113) | -0.0613*** (0.0113) |
| Lender Log GDP | 0.210*** (0.0156) | 0.209*** (0.0156) | 0.0955*** (0.0167) | 0.0952*** (0.0169) | 0.0723*** (0.0238) | 0.0711*** (0.0238) | 0.0464** (0.0233) | 0.0465** (0.0233) |
| Borrower Log GDP | 0.123*** (0.0118) | 0.124*** (0.0118) | 0.0656*** (0.0135) | 0.0656*** (0.0135) | 0.0307 (0.0187) | 0.0309* (0.0187) | 0.0413** (0.0205) | 0.0408** (0.0204) |
| Language * Region | -0.0261 (0.0207) | -0.0429** (0.0199) | -0.0481*** (0.0161) | -0.0555*** (0.0152) | -0.0490** (0.0209) | -0.0664*** (0.0199) | -0.0363* (0.0187) | -0.0447** (0.0178) |
| Colonizer * Region | -0.0429 (0.0298) | -0.0898*** (0.0265) | -0.00595 (0.0598) | -0.0510 (0.0432) | -0.0644** (0.0296) | -0.109*** (0.0250) | -0.0168 (0.0600) | -0.0596 (0.0400) |
| Legal * Region | 0.0209 (0.0143) | 0.0132 (0.0138) | -0.00706 (0.0127) | -0.0127 (0.0124) | 0.0273* (0.0160) | 0.0179 (0.0155) | -0.0156 (0.0135) | -0.0220* (0.0130) |
| Colonial Relation * Region | -0.0581 (0.0363) | -0.0552 (0.0356) | -0.0437 (0.0280) | -0.0388 (0.0289) | -0.0233 (0.0431) | -0.0191 (0.0423) | -0.0527** (0.0250) | -0.0477* (0.0264) |
| FTA * Region | -0.0520*** (0.0169) | -0.0304* (0.0176) | -0.0353* (0.0186) | -0.0297 (0.0183) | -0.0368* (0.0199) | -0.0131 (0.0207) | -0.0285 (0.0208) | -0.0214 (0.0205) |
| Log Distance * Region | 0.0967*** (0.0126) | 0.0784*** (0.0129) | 0.0253* (0.0132) | 0.0186 (0.0130) | 0.102*** (0.0136) | 0.0809*** (0.0140) | 0.0337** (0.0138) | 0.0251* (0.0136) |
| Concentration | -0.0479** (0.0234) | -0.0496** (0.0233) | 0.0271 (0.0226) | 0.0279 (0.0226) | -0.0524* (0.0302) | -0.0556* (0.0301) | 0.00358 (0.0274) | 0.00304 (0.0274) |
| Overhead | -0.0537*** (0.0109) | -0.0588*** (0.0109) | -0.0169 (0.0137) | -0.0242* (0.0138) | -0.0780*** (0.0126) | -0.0833*** (0.0127) | -0.0128 (0.0141) | -0.0211 (0.0144) |
| Entry Cost | | | -0.0329* (0.0175) | -0.0373** (0.0180) | | | -0.0367* (0.0192) | -0.0420** (0.0198) |
| Entry Time | | | -0.00619 (0.00981) | -0.00858 (0.0101) | | | 0.00297 (0.0122) | 0.00108 (0.0124) |
| N | 139216 | 139216 | 91996 | 91996 | 83362 | 83362 | 53194 | 53194 |
| Pseudo R-sq | 0.229 | 0.231 | 0.264 | 0.267 | 0.192 | 0.194 | 0.230 | 0.233 |
| Lender + Borrower FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES | YES | YES | YES | YES |

Notes: This table reports the first-stage Probit regression result on the baseline model, with the common region indicator serving as a gravity factor. Its interaction with traditional, static gravity factors are also included. Standard errors are clustered at country-pair level. All coefficients are marginal effects. "Post-crisis" refers to the period 2009Q3-2014Q4.

We use the same framework to answer this question, by including a set of interactions between the regional dummy and other *static* gravity factors in our baseline estimations. The results are presented in Table 7 and 8. In many cases, regional proximity seems to have partially reduced the barriers to cross-border lending, reflected by the negative signs of its first-stage interactions with common language and common colonizer dummy, as well as the positive

Table 8. Region as an Interaction Term – Second Stage

| | Full Sample | | | | Post-Crisis Sample | | | |
|----------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | Direct Cross-Border | | Local Affiliate | | Direct Cross-Border | | Local Affiliate | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Region | -1.644 (1.126) | -0.587 (1.183) | 2.422 (2.956) | 2.115 (2.771) | -1.783 (1.630) | -0.570 (1.463) | 4.423 (5.075) | 2.850 (4.099) |
| Region * Peripheral | | 0.927*** (0.188) | | 0.781 (0.806) | | 0.736*** (0.260) | | 0.0347 (1.472) |
| Language | 0.177 (0.112) | 0.220* (0.115) | 0.609 (0.375) | 0.675* (0.391) | 0.261* (0.142) | 0.293** (0.144) | 0.368 (0.650) | 0.501 (0.635) |
| Colonizer | 0.675*** (0.179) | 0.827*** (0.187) | 2.724*** (0.882) | 2.931*** (0.902) | 0.605*** (0.223) | 0.704*** (0.235) | 2.935* (1.569) | 3.034* (1.600) |
| Legal | 0.0688 (0.0653) | 0.0620 (0.0665) | 0.0198 (0.273) | 0.0677 (0.274) | 0.0758 (0.0777) | 0.0722 (0.0783) | -0.428 (0.481) | -0.304 (0.450) |
| Colonial Relation | 0.995*** (0.130) | 0.888*** (0.141) | 1.456*** (0.459) | 1.430*** (0.421) | 0.794*** (0.200) | 0.666*** (0.190) | 1.024 (1.071) | 1.230 (0.856) |
| FTA | 0.137 (0.0871) | 0.171* (0.0875) | 0.136 (0.314) | 0.206 (0.333) | 0.0176 (0.109) | 0.0505 (0.110) | -0.146 (0.604) | -0.0464 (0.617) |
| Log Distance | -1.135*** (0.131) | -0.964*** (0.146) | -1.113** (0.524) | -1.157** (0.477) | -1.111*** (0.220) | -0.932*** (0.196) | -0.431 (1.093) | -0.724 (0.861) |
| Lender Log GDP | 0.868*** (0.115) | 0.691*** (0.166) | -0.345 (0.366) | -0.263 (0.351) | 0.588*** (0.192) | 0.450** (0.178) | -0.388 (0.602) | -0.240 (0.546) |
| Borrower Log GDP | 0.724*** (0.0891) | 0.631*** (0.114) | 1.246*** (0.394) | 1.223*** (0.375) | 0.607*** (0.145) | 0.551*** (0.141) | 0.865 (0.648) | 0.970* (0.584) |
| Language * Region | -0.0911 (0.191) | -0.221 (0.189) | -0.766 (0.621) | -0.834 (0.639) | -0.0936 (0.230) | -0.188 (0.236) | -0.381 (0.908) | -0.455 (0.924) |
| Colonizer * Region | -0.382 (0.247) | -0.789*** (0.269) | -2.303** (1.039) | -2.820*** (1.087) | -0.176 (0.303) | -0.454 (0.342) | -2.382 (1.698) | -2.469 (1.770) |
| Legal * Region | 0.340*** (0.121) | 0.287** (0.120) | 0.615* (0.350) | 0.576 (0.357) | 0.382*** (0.146) | 0.323** (0.142) | 0.924** (0.458) | 0.860* (0.487) |
| Colonial Relation * Region | -0.638* (0.365) | -0.492 (0.356) | -0.171 (0.714) | -0.132 (0.680) | -0.630 (0.415) | -0.479 (0.393) | -0.200 (1.293) | -0.382 (1.084) |
| FTA * Region | -0.0185 (0.146) | 0.123 (0.143) | -0.0172 (0.506) | -0.0197 (0.501) | 0.209 (0.171) | 0.300* (0.162) | 0.115 (0.752) | 0.0176 (0.654) |
| Log Distance * Region | 0.179 (0.131) | 0.00719 (0.135) | -0.309 (0.358) | -0.302 (0.319) | 0.177 (0.191) | -0.00501 (0.164) | -0.557 (0.648) | -0.360 (0.471) |
| δ | 0.209* (0.109) | 0.505 (0.526) | 4.632*** (1.566) | 4.388*** (1.525) | 2.204*** (0.818) | 2.607*** (0.781) | 5.508* (2.845) | 4.780* (2.449) |
| β | 1.025*** (0.235) | 0.865*** (0.237) | 0.923 (0.932) | 0.912 (0.872) | 1.270*** (0.452) | 1.115*** (0.420) | -1.001 (2.372) | -0.540 (1.969) |
| σ | 1.533*** (0.0121) | 1.528*** (0.0121) | 2.304*** (0.0393) | 2.301*** (0.0393) | 1.550*** (0.0149) | 1.546*** (0.0149) | 2.402*** (0.0485) | 2.399*** (0.0486) |
| N | 37625 | 37625 | 11044 | 11044 | 20221 | 20221 | 5974 | 5974 |
| Lender + Borrower FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES | YES | YES | YES | YES |

Notes: This table reports the second-stage maximum likelihood estimation result on the baseline model, with the common region indicator serving as a gravity factor. Its interaction with traditional, static gravity factors are also included. Standard errors are clustered at country-pair level. δ and β refer to, respectively, the parameter of the non-linear adjustment term and the estimated inverse Mill's ratio term in Equation (20). σ is the estimated standard deviation of the disturbance term, e_{ij} , in Equation (20). "Post-crisis" refers to the period 2009Q3-2014Q4.

sign of its first-stage interaction with geographical distance. On the other hand, positive synergies are present between regional affiliations and common legal system at the intensive margin. We find that the barrier imposed by distance for both direct cross-border and local affiliate banking is significantly reduced by sharing the same region, even if the geographical distance between the lender-borrower pair in the same region (for instance, the continent of the Americas) could be large.

The interpretation of the regional dummy changes accordingly once we include its interactions with other gravity factors. At the extensive margin, we find a distinct, significant, yet negative regional effect in determining global banking flows, especially in the post-crisis subset. However, as we further interact the regional dummy with the the group indicator of peripheral lenders, we find that the interaction terms come with strongly significant and positive coefficients, implying that the negative regional effect on its own is more likely due to the reshuffling of regional banking systems, such as the Euro Area, dominated by traditional global banks and their affiliates. The finding is also clear at the intensive margin for direct cross-border flows, as the coefficient of the interaction between common region and peripheral lenders is of the opposite sign and significantly different from the coefficient of the regional dummy. In the context, by accounting for possible synergies with other gravitational factors, we have found that regional proximity seems to dampen the barriers captured by geographical distance to direct cross-border and local affiliate lending at the extensive margin, as well as increasing the role of same legal origin at the intensive margin. Moreover, after accounting for those synergies, the regional dummy alone seems to also capture unique determinants of global banking, with the peripheral lenders playing a particularly different role, as we have found in the previous section.

VII. WHAT ELSE IS THE REGIONAL DUMMY CAPTURING?

In the previous two sections, we have established that regionalization led by peripheral lenders is an increasingly important characteristic of global banking network, which is present even after controlling for gravitational variables. As the natural next step, we focus on three major areas of post-crisis transformation in global banking that might characterize structural changes and regional fragmentation.

Regulatory Environment. Past literature has explored the close connection between international banking flows and regulation. [Houston, Lin, and Ma \(2012\)](#) use 1996-2007 BIS CBS data and find evidence that banking flows have shifted towards countries with lighter regulations. After the crisis, however, with the overall trend of tightening regulation, the space for regulatory arbitrage is depressed, so that the interaction between banking flows and regulation may become unclear. On the other hand, facing increasing restrictions, banks may seek relative information advantage, preferring to lend to areas with common ties.

Trade Linkages. The interaction between real linkages and financial linkages has been extensively studied in an international context. Past research has identified a positive relationship

between bilateral trade and cross-border exposures (Blank and Buch, 2010; Cerutti, Hale, and Minoiu, 2015). Like global financial integration, international trade has also suffered a change following the financial crisis. Still, trade finance remains an integral part of potential drivers that promotes regional financial integration.

Global Banks' Retrenchment. Following the global financial crisis, global banks from core and European lenders have been divesting from emerging markets. Teodoru (2017) identifies forty-two cases of large, global banking groups selling operations or assets in Latin America, Middle East, Africa, Emerging Europe and Emerging Asia since 2009. Such a retrenching pattern, while raising concerns on the potential underfinancing of emerging market borrowers, comes with an opportunity for the peripheral lenders to utilize existing banking resources and develop regional lending in the market originally dominated by global banks.

To test if the above factors influence the regionalization process, we obtain data on bank ownership, banking regulations and systemic banking crises, and create additional interactions with common region indicators in our empirical model. Table 9 contains the definitions and sources of our variables for interaction, and Table 10 reports the summary statistics. We use lender's export to borrower as share of borrower's total import to proxy for the relative strength of trade linkages between the country pair. For regulation variables, we utilize the database compiled by Barth, Caprio, Jr, and Levine (2013) from four rounds of surveys completed by bank regulatory officials from 118 countries.¹⁸ In particular, we choose the variable "overall activity restriction", as it comes with better coverage and is supposed to serve as a good proxy for the restrictiveness of banking regulations. In addition to using only the borrowers' regulatory level, we calculate the bilateral regulatory distance between each pair of lender and borrower to supplement the analysis.

For retrenchment variables, we restrict to the sample containing only the linkages from peripheral lenders. For each borrower, we calculate the share of borrowing from core and European lenders that, over the sample period, have experienced systemic banking crises, given by Laeven and Valencia (2012).¹⁹ Meanwhile, using the comprehensive bank ownership database by Claessens and van Horen (2015), we calculate the total number of banks operating in each borrower country and owned by lenders that have experienced systemic banking crises. We expect as traditional global banks under stress revamp their business models and scale down their global footprints, captured by a decrease in the share of exposure and in the

¹⁸We use the latest three rounds of survey, and assign each round of survey to one crisis episode. Survey 2 (2003) is used to proxy for pre-crisis regulation environment, while Survey 3 (2007) and Survey 4 (2011) represent crisis and post-crisis bank regulation levels, respectively.

¹⁹Countries include Austria, Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Switzerland, United Kingdom and United States.

number of subsidiaries owned, peripheral lenders fill in the vacancy by forming linkages and increasing their claims to countries where global banks retreat.²⁰ We transform these two measures by taking differences between lender and borrower's shares over a specified time horizon, to better capture the notion of retrenchment.²¹ Finally, we construct triple interactions of the differenced measures with common region, post-crisis and crisis dummies, and feed them one by one into our original model.

Table 9. Variable Definitions: Interaction Variables

| Variable | Definition | Source |
|-----------------------|--|--|
| Overall Restriction | The extent to which banks may engage in securities, insurance and real estate activities. Higher value indicates higher restrictiveness. | Barth, Caprio, Jr, and Levine (2013) |
| Trade Linkage | Lender's export to borrower as share of total imports of borrower. | Direction of Trade Statistics |
| Share of Exposure | Core and European lenders in systemic banking crises: exposure to borrower as share of borrower's total exposure. | Author's calculation, Laeven and Valencia (2012) |
| Foreign Bank Presence | Core and European lenders in systemic banking crises: number of banks owned in borrower country. | Author's calculation, Claessens and van Horen (2015); Laeven and Valencia (2012) |

Notes: For regulation variables (overall restriction, conglomerate restriction, statement transparency), three levels of values are assigned to each time period associated with the crisis. Survey 2 (2003) is associated with pre-crisis (2006Q1-2007Q4) regulation level; Survey 3 (2007) is assigned to crisis-time (2008Q1-2009Q2) regulation level; Survey 4 (2011) is used to proxy post-crisis (2009Q2 onwards) regulation. For retrenchment variables (share of exposure and foreign bank presence), coefficients of the interaction terms are estimated only on the restricted sample where only the connections between peripheral lenders and all borrowers are preserved. Trade variables used in the estimation are lagged by one quarter. "Foreign bank presence" at year t is further transformed by taking the difference between year $t - 1$ and year $t - 2$, while "share of exposure" at quarter i is transformed by four-quarter difference (quarter i and quarter $i - 4$).

In the data, peripheral lenders and non-reporting countries put relatively more restrictions on bank activities. As borrower countries increase restrictions on the activity of banks, the gap between lender and borrower in bank regulation becomes more negative. Therefore, when comparing results between borrower and lender-borrower variables, we should expect the coefficients of two measures to be of opposite sign if the results are consistent with each other. Indeed, for both direct cross-border and local affiliate flows, the overall interaction of the

²⁰The bank ownership database only covers data prior to 2014. As a result, recent major transactions, such as Scotiabank's 2015 acquisition of Citigroup's retail and commercial banking operations in Panama and Costa Rica is not included in the data. Including the most recent cases will further bolster our claim.

²¹We assign the change in the number of banks owned by countries in stress as follows: for year t , the number is equal to the change from year $t - 2$ to year $t - 1$, as we acknowledge set-up costs and time for regulatory approval. For the change in exposure, we take 4Q differences from the current quarter.

Table 10. Summary Statistics: Additional Interaction Variables

| Sample | Variable | N | Mean | SD | Min | Max |
|--------------------------------------|--------------------------------|--------|-------|-------|-----|-------|
| All Countries | Lenders: Overall Restriction | 83 | 6.49 | 1.94 | 3 | 10 |
| | Borrowers: Overall Restriction | 396 | 7.48 | 1.98 | 3 | 12 |
| | Trade Linkages | 210420 | 1.93 | 4.92 | 0 | 90.94 |
| Peripheral Lenders and All Borrowers | Share of CB Exposure | 8077 | 75.98 | 26.71 | 0 | 100 |
| | Share of LC Exposure | 4666 | 68.64 | 39.44 | 0 | 100 |
| | Foreign Bank Presence | 1746 | 3.72 | 6.84 | 0 | 60 |

Notes: Summary statistics presented in this table are computed from the raw data before further differencing and transformation. For retrenchment variables (share of exposure, foreign bank presence, degree of connection), CB exposure (connection) is used whenever the dependent variable is cross-border, vice versa for LC exposure (connection). *N* in the table denotes the number of unique observations at country level or country-pair level.

regional dummy and borrower regulation level is significantly positive (Table 11, specification 1, 5), while the overall interaction of the regional dummy and lender-borrower regulation distance is significantly negative (Table 11, specification 3, 7) at the first stage. As we further decompose the coefficients by adding interactions with post-crisis and crisis dummies, we find that the interaction between common region and regulatory constraints seems to provide additional explanatory power, mostly during the crisis and post-crisis periods. For both types of flows, we find a positive and significant second-stage coefficient for the base level interaction in Table 11, specification 4 and 8. Consistent with the analysis of [Houston, Lin, and Ma \(2012\)](#) using a pre-crisis sample, these findings imply an increase in direct cross-border and local affiliate outflow, following a increase in bilateral regulatory distance. However, the finding of a prevalent regulatory arbitrage in [Houston, Lin, and Ma \(2012\)](#) seems to no longer apply to the post-crisis period, as the positive interaction coefficients for both direct cross-border and local affiliate flows (Table 11, specification 2 and 6, as well as the negative coefficients with the lender-borrower variables in specification 4 and 8) suggest the opposite message after the crisis, controlling for other factors at the first stage. As the regulation gap between lenders and borrowers narrows after the crisis, more within-region connections are established for both direct cross-border and local affiliate lending at the extensive margin, and direct cross-border flows also grow in intensity. According to [Barth, Caprio, Jr, and Levine \(2013\)](#), a large proportion of countries tightened regulations after the great financial crisis, as reflected in the 2011 survey. Instead of moving funds to lightly regulated counterparts, regional banks seem to be propelled to adapt to the shifting global policy environment after the crisis. In this way, we echo [Karolyi and Taboada \(2015\)](#), who find that cross-border bank acquisitions do not necessarily correspond with a potentially destructive “race to the bottom” by targeting countries with the weakest regulatory framework.

Table 11. Interaction with Bank Activity Restrictions

| Direct Cross-Border First Stage | Borrower | | Lender – Borrower | |
|--|------------------------|------------------------|--------------------------|-------------------------|
| | (1) | (2) | (3) | (4) |
| Region | 0.0204*** (0.00308) | -0.00792* (0.00433) | -0.00935*** (0.00235) | 0.000680 (0.00335) |
| Region * Crisis | | 0.0407*** (0.00590) | | -0.00757 (0.00481) |
| Region * Post-crisis | | 0.0384*** (0.00551) | | -0.0149*** (0.00439) |
| N | 104020 | 104020 | 96552 | 96552 |
| Pseudo R-sq | 0.212 | 0.214 | 0.212 | 0.214 |
| Second Stage | | | | |
| Region | 0.0293 (0.0254) | 0.0201 (0.0269) | 0.0180 (0.0174) | 0.0547*** (0.0203) |
| Region * Crisis | | 0.0648 (0.0398) | | -0.0512* (0.0277) |
| Region * Post-crisis | | -0.00656 (0.0428) | | -0.0608** (0.0289) |
| N | 30941 | 30941 | 28732 | 28732 |
| Local Affiliate First Stage | | | | |
| | Borrower | | Lender – Borrower | |
| | (5) | (6) | (7) | (8) |
| Region | 0.0109*** (0.00285) | 0.00699* (0.00395) | -0.00754*** (0.00227) | -0.00102 (0.00330) |
| Region * Crisis | | 0.0102** (0.00464) | | -0.00735** (0.00367) |
| Region * Post-crisis | | 0.00375 (0.00489) | | -0.00969** (0.00418) |
| N | 73880 | 73880 | 67232 | 67232 |
| Pseudo R-sq | 0.251 | 0.251 | 0.259 | 0.259 |
| Second Stage | | | | |
| Region | 0.0239 (0.0754) | -0.0448 (0.0817) | 0.108** (0.0519) | 0.124** (0.0568) |
| Region * Crisis | | 0.0257 (0.0875) | | 0.0792 (0.0648) |
| Region * Post-crisis | | 0.131 (0.0985) | | -0.0799 (0.0897) |
| N | 9700 | 9700 | 8901 | 8901 |

Notes: All coefficients for the first stage are average marginal effects. Standard errors are clustered at country-pair level. Each cell records the estimated coefficient of the interaction between the column variable and the row variable. For example, the coefficient of specification (1) refers to the coefficient of the interaction term “Region * Overall Restriction”. “Borrower” refers to the overall activity restriction index of the borrower country, while “lender – borrower” refers to the difference in overall activity restriction index between the lender and borrower pair.

The additional contribution of trade linkages to regional banking flows, controlling for traditional gravity factors, is small albeit significant, as is shown in Table 12. Overall, the formation of the banking linkages does not seem to strongly synchronize with trade relationships, as is reflected in negative average marginal effects at the first stage (specification 1 and 3). The crisis does not significantly alter the relationship, and both the marginal effect and the estimated coefficients at the intensive margin are weak and close to zero.

Table 12. Interaction with Trade Linkages

| First Stage | Direct Cross-Border | | Local Affiliate | |
|-------------------------------------|----------------------------|---------------------------|---------------------------|--------------------------|
| | (1) | (2) | (3) | (4) |
| Region * Export Share | -0.00186** (0.000858) | -0.00326*** (0.000968) | -0.00202*** (0.000682) | -0.00189** (0.000856) |
| Region * Crisis * Export Share | | 0.00120 (0.00117) | | -0.000561 (0.000621) |
| Region * Post-crisis * Export Share | | 0.00196** (0.000968) | | -0.0000890 (0.000709) |
| N | 135932 | 135932 | 91184 | 91184 |
| Pseudo R-sq | 0.220 | 0.222 | 0.260 | 0.260 |
| Second Stage | | | | |
| Region * Export Share | -0.00723 (0.00659) | -0.0141* (0.00789) | 0.0261 (0.0206) | 0.0214 (0.0255) |
| Region * Crisis * Export Share | | 0.00615 (0.00572) | | 0.0106 (0.0144) |
| Region * Post-crisis * Export Share | | 0.0112* (0.00598) | | 0.00647 (0.0182) |
| N | 36870 | 36870 | 10763 | 10763 |

Notes: This table reports two-stage results on the interaction of common region, crisis episode variables with trade linkages, defined as lender's bilateral export to borrower as share of borrower's total imports. For first-stage results, coefficients are average marginal effects. At both stages, standard errors are clustered at country-pair level. Trade variables used in the estimation are lagged by one quarter.

Finally, Table 13 displays the regression results for the interactions with retrenchment variables, with the estimations implemented on the restricted sample. Anecdotal evidence on advanced countries' post-crisis spin-off of bank operations in emerging markets (Teodoru, 2017) anticipates a substitution by regional flows from peripheral lenders, in order to compensate for the loss of credit originally provided by global banks. This rising importance of peripheral lenders in their regional markets is indeed heavily supported by our regressions, especially when the dependent variable is local affiliate connection. First-stage regressions, in particular, forms a unified picture of peripheral lenders' response to global banks' retrenchment. In Table 13, the overall interaction between the regional dummy and advanced, stressed lenders' change in direct cross-border exposure is close to zero and insignificant (specification 7). As soon as we further interact with the post-crisis dummy, the estimated marginal

Table 13. Peripheral Lenders: Interaction with Global Banks' Retrenchment

| Direct Cross-Border First Stage | Δ Share of CB Exposure | | Δ Share of LA Exposure | | Δ Bank Presence | |
|--|-------------------------------|---------------------------|-------------------------------|---------------------------|------------------------|------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Region | -0.00125*** (0.000368) | -0.00115* (0.000641) | -0.000537** (0.000261) | -0.000392 (0.000496) | -0.0150** (0.00709) | 0.00974 (0.0179) |
| Region * Crisis | | 0.000264 (0.000857) | | -0.000403 (0.000837) | | -0.0446** (0.0226) |
| Region * Post-crisis | | -0.000460 (0.000841) | | -0.000172 (0.000609) | | -0.0252 (0.0216) |
| N | 48596 | 48596 | 29255 | 29255 | 48756 | 48756 |
| Pseudo R-sq | 0.301 | 0.301 | 0.242 | 0.242 | 0.302 | 0.303 |
| Second Stage | | | | | | |
| Region | -0.00756* (0.00430) | -0.00201 (0.00624) | 0.00289 (0.00185) | 0.00418 (0.00277) | -0.0685 (0.0670) | -0.143 (0.110) |
| Region * Crisis | | -0.00284 (0.0134) | | -0.000912 (0.00641) | | 0.242 (0.149) |
| Region * Post-crisis | | -0.00940 (0.00785) | | -0.00192 (0.00360) | | -0.0229 (0.147) |
| N | 10367 | 10367 | 8274 | 8274 | 10367 | 10367 |
| Local Affiliate First Stage | | | | | | |
| | Δ Share of CB Exposure | | Δ Share of LA Exposure | | Δ Bank Presence | |
| | (7) | (8) | (9) | (10) | (11) | (12) |
| Region | -0.000208 (0.000643) | 0.00111** (0.000515) | -0.000615*** (0.000236) | -0.00132*** (0.000503) | -0.00716 (0.00793) | 0.0313* (0.0167) |
| Region * Crisis | | -0.00154 (0.00210) | | 0.000648 (0.000766) | | -0.0370** (0.0155) |
| Region * Post-crisis | | -0.00280*** (0.000960) | | 0.000932 (0.000571) | | -0.0520*** (0.0195) |
| N | 18540 | 18540 | 16281 | 16281 | 18540 | 18540 |
| Pseudo R-sq | 0.279 | 0.280 | 0.268 | 0.268 | 0.279 | 0.280 |
| Second Stage | | | | | | |
| Region | -0.00969 (0.00952) | -0.00710 (0.0194) | -0.00539 (0.0101) | -0.0113 (0.0166) | -0.00633 (0.179) | 0.383 (1.267) |
| Region * Crisis | | 0.0482 (0.0334) | | 0.00599 (0.0120) | | -0.405 (1.113) |
| Region * Post-crisis | | 0.00558 (0.0349) | | 0.00390 (0.0158) | | -0.287 (1.321) |
| N | 2292 | 2292 | 2254 | 2254 | 2292 | 2292 |

Notes: Estimations are performed on the restricted sample where only connections between peripheral lenders and all borrowers are preserved. All coefficients are average marginal effects for the first stage. Standard errors are clustered at country-pair level. Each cell correspond to the interaction between the column variable and the row variable. For example, the coefficient of specification (1) refers to the coefficient of the interaction term "Region * Δ Share of Cross-Border Exposure". "Bank presence" at year t is transformed by taking the difference between year $t - 1$ and year $t - 2$, while "share of exposure" at quarter i is transformed by four-quarter difference (quarter i and quarter $i - 4$).

effect becomes highly significant and negative (specification 8). Through an increase in local affiliate exposures, peripheral lenders alleviate the concern of their regional counterparts that have lost cross-border credit due to tightened balance sheets of global banks. To increase their local affiliate claims, peripheral lenders need to enter the market by opening branches or acquiring subsidiaries. The interaction with foreign bank presence captures this strategy. Comparing the first-stage results of specification 11 and 12, the triple interaction adding the post-crisis dummy is again strongly significant and negative, and the estimated marginal effect is large. A decrease in the number of banks owned by advanced, stressed lenders reflects ownership changes, as the operations are typically sold to regional, peripheral lenders. In addition, peripheral lenders can also increase their market share in a target country by acquiring a domestically owned bank. Utilizing the existing customer base and banking resources while holding an informational and cultural advantage compared to core and European countries, peripheral lenders are able to expand their local affiliate claims to borrowers through acquired branches and subsidiaries within the same region.²²

While we do not find comparably significant post-crisis interaction terms for direct cross-border flows, the negative coefficient of the overall interaction between the regional dummy and the retrenchment proxies are significant across specification 1, 3, and 5, indicating that for direct cross-border flows, substitution across lenders seems to have existed before the crisis. The second stage regressions for both types of flows yield insignificant results, suggesting that peripheral lenders may not have yet matched the level of funding from their local affiliates when compared with their global, advanced predecessors.

VIII. CONCLUSION

There was a shrinkage in the overall amount of cross-border bank lending after the GFC. Nonetheless, these dynamics do not follow a generalized retrenchment in cross-border bank lending since some parts of the banking network are now more regionally linked than before the GFC. In this paper, using a two-step strategy, we analyzed in detail the nature of this regionalization phenomenon. First, taking into account that the regionalization process is

²²While global banks often sell their operations directly to the countries in which the operations are based, such as Deutsche Bank's 2016 agreement with Banco Comafi in Argentina to sell its Argentine bank subsidiary, there exist many cases for which the buyers are different. Peripheral lenders may further consolidate their positions in the local market by merging their new acquisitions from global banks and domestic banks. According to official releases and the data of [Claessens and van Horen \(2015\)](#), Canadian multinational bank group Scotiabank purchased the shares of Royal Bank of Scotland (Colombia) as Scotiabank Colombia, and merged its operation with Banco Colpatria, the fifth largest bank of Colombia, after it acquired the latter in 2013.

happening especially in the periphery of the banking network (where linkages are far from complete) and that foreign banks' preference for regional cross-border lending could be capturing multidimensional factors (e.g., from traditional geographical and cultural factors to institutional and idiosyncratic factors after GFC), we developed a gravitational model of banks' cross-border lending expansion, combining recent advances in both trade (Helpman, Melitz, and Rubinstein, 2008) and banking literature (Buch, Koch, and Koetter, 2011, 2014; Fillat and others, 2017). More specifically, our theoretical framework allowed us to investigate the regionalization phenomenon at a global level, explicitly connecting banks' lending / no-lending decision with the quantity-of-lending decision, and to obtain consistent estimates using only aggregate country-level bilateral data, as in the case of trade. Second, we estimated the gravitational model using the BIS Consolidated Banking Statistics (the most comprehensive global dataset available that allows a breakdown between direct cross-border and local affiliate lending) to which, following Cerutti (2015) and Cerutti and Zhou (2017), we performed some adjustments in order to depict a more accurate picture of bilateral banking linkages (e.g., excluding local deposit funding from local affiliate claims).

The estimation results indicate that the regionalization trend was present (in both direct cross-border and local affiliate lending) before the GFC, but it has increased since then, and it is being driven by banks in EMs and non-core banking systems, such as Australia, Canada, Hong Kong, and Singapore. Moreover, the regionalization trend is capturing factors linked to both standard gravitational variables and institutional and idiosyncratic factors. With respect to gravitational factors, capturing information asymmetries, we find that the regionalization trend seems to be associated with the dampening of some barriers to direct cross-border and local affiliate lending (captured by geographical distance) at the extensive margin, as well as increasing the role of others (same legal origin) at the intensive margin. Even more interestingly, the regionalization phenomenon seems to reflect additional unique dimensions, especially at the extensive margin. Regional banks seem to prefer increasing their operations toward countries with stricter regulations within the same region; and some of the increase in the regionalization trends by non-core lenders seems to be associated with taking advantage of the opportunities created by the retrenchment of several European lenders, both in terms of direct cross-border and affiliate local claims. Finally, given the nature of the BIS sample used that does not include China and South Korea, and other potential small recent regional lenders as Colombia, our findings of the presence of regionalization in both direct cross-border and affiliates' local lending is reassuring since the sample used in our calculations is biased against this phenomenon.

The fact that the increasing regionalization trend is not only driven by some one-off but likely persistent factors (e.g. retrenchment of European banks due to crisis), and also by regional advantages vis-à-vis traditional gravitational factors as well as regulation implies that its higher importance after the GFC is most likely not a transitory phenomenon. As in the case of global foreign banks, the larger presence of regional foreign banks can have both positive and negative aspects. In principle, on the one hand, it is possible that an increase in regional coordination and supervision collaboration could be more feasible to achieve (e.g. European experience through various initiatives such as Vienna Initiative). This is especially the case after the GFC, with some new regulations having moved away from global initiatives (e.g., recent geographical ring-fencing in the UK, and 2014 US foreign bank organization rules). [Claessens and van Horen \(2016\)](#) finds that the presence of foreign banks tend to be accompanied by credit booms in the host country before the crisis. A regionalized structure after the crisis, therefore, may alleviate the concern of local busts as regional banks may come with small scale of operation. On the other hand, the regional trend could trigger more fragmentation with adverse consequences for risk-sharing and financial stability. For example, many new regional lenders do not have a long tradition of monitoring the cross-border activities of their banks. The funding structure of regional EM banks is more limited than global banks and more sensitive to regional shocks. These limitations and others could trigger local authorities in host countries to encourage more restrictive regulation (e.g., higher local funding requirements, tougher restrictions on activities, higher capital, reserve, and provisioning requirements directed to insulate foreign affiliates from their parent banks as documented in [Anginer, Cerutti, and Martínez Pería \(2017\)](#)). In this context, as our results seems to indicate, the changes in the post-crisis regulatory frameworks might not be conducive to a very diverse or inclusive environment. More research is needed to analyze the nature and potential consequences of the differences between global and regional cross-border banking relationships after the GFC.

REFERENCES

- Aiyar, Shekhar, Charles W. Calomiris, John Hooley, Yevgeniya Korniyenko, and Tomasz Wieladek, 2014, “The international transmission of bank capital requirements: Evidence from the UK,” *Journal of Financial Economics*, Vol. 113, pp. 368–382.
- Anderson, James E., and Yoto V. Yotov, 2017, “Short Run Gravity,” NBER Working Paper 23458.
- Anginer, Deniz, Eugenio Cerutti, and María Soledad Martínez Pería, 2017, “Foreign bank subsidiaries’ default risk during the global crisis: What factors help insulate affiliates from their parents?” *Journal of Financial Intermediation*, Vol. 29, pp. 19–31.
- Bank for International Settlements, 2014, “EME Banking systems and regional financial integration,” Committee on the Global Financial System Paper 51.
- Barth, James R., Gerard Caprio, Jr, and Ross Levine, 2013, “Bank Regulation and supervision in 180 countries from 1990 to 2011,” *Journal of Financial Economic Policy*, Vol. 5, No. 2, pp. 111–219.
- Beck, Thorsten, Asli Demirguc-Kunt, and Ross Levine, 2000, “A New Database on Financial Development and Structure,” *World Bank Economic Review*, Vol. 14, pp. 597–605.
- , 2009, “Financial Institutions and Markets Across Countries and over Time: Data and Analysis,” World Bank Policy Research Working Paper 4943.
- Blank, Sven, and Claudia M. Buch, 2010, “International bank portfolios: Short- and Long-Run Responses to Macroeconomic Conditions,” *Review of International Economics*, Vol. 18, No. 2, pp. 289–306.
- Bouvatier, Vincent, and Anne-Laure Delatte, 2015, “Waves of international banking integration: A tale of regional differences,” *European Economic Review*, Vol. 80, pp. 354–373.
- Brei, Michael, and Goetz von Peter, 2017, “The distance effect in banking and trade,” BIS Working Papers No. 658.
- Bremus, Franziska, Claudia M. Buch, Katheryn N. Russ, and Monika Schnitzer, 2014, “Big banks and macroeconomic outcomes: Theory and cross-country evidence of granularity,” Working Paper.
- Buch, Claudia M., Cathérine T. Koch, and Michael Koetter, 2011, “Size, productivity, and international banking,” *Journal of International Economics*, Vol. 85, pp. 329–334.
- , 2014, “Should I stay or should I go? Bank productivity and internationalization decisions,” *Journal of Banking and Finance*, Vol. 42, pp. 266–282.
- Buch, Claudia M., and Alexander Lipponer, 2007, “FDI versus exports: Evidence from German banks,” *Journal of Banking and Finance*, Vol. 31, pp. 805–826.

- Cerutti, Eugenio, 2015, “Drivers of cross-border banking exposures during the crisis,” *Journal of Banking and Finance*, Vol. 55, pp. 340–357.
- Cerutti, Eugenio, and Stijn Claessens, 2017, “The great cross-border bank deleveraging: Supply constraints and intra-group frictions,” *Review of Finance*, Vol. 21, No. 1, pp. 201–236.
- Cerutti, Eugenio, Giovanni Dell’Ariccia, and Maria Soledad Martinez Peria, 2007, “How banks go abroad: Branches or subsidiaries?” *Journal of Banking and Finance*, Vol. 31, No. 6, pp. 1669–1692.
- Cerutti, Eugenio, Galina Hale, and Camelia Minoiu, 2015, “Financial crises and the composition of cross-border lending,” *Journal of International Money and Finance*, Vol. 52, pp. 60–81.
- Cerutti, Eugenio, and Haonan Zhou, 2017, “The Global Banking Network in the Aftermath of the Crisis: Is There Evidence of De-globalization?” IMF Working Paper 17/232.
- Cetorelli, Nicola, and Linda S. Goldberg, 2012, “Banking Globalization and Monetary Transmission,” *The Journal of Finance*, Vol. 67, No. 5, pp. 1811–1843.
- Cihak, Martin, Asli Demirguc-Kunt, Erik Feyen, and Ross Levine, 2012, “Benchmarking Financial Development Around the World,” World Bank Policy Research Working Paper 6175.
- Claessens, Stijn, and Neeltje van Horen, 2012, “Being a foreigner among domestic banks: Asset or liability?” *Journal of Banking and Finance*, Vol. 36, No. 5, pp. 1276–1290.
- , 2014a, “Foreign banks: Trends and Impact,” *Journal of Money, Credit and Banking*, Vol. 46, No. s1, pp. 295–326.
- , 2014b, “Location decisions of foreign banks and competitor remoteness,” *Journal of Money, Credit and Banking*, Vol. 46, No. 1, pp. 145–170.
- , 2015, “The impact of the global financial crisis on banking globalization,” *IMF Economic Review*, Vol. 63, No. 4, pp. 868–918.
- , 2016, “The Role of Foreign Banks in Local Credit Booms,” in *World Scientific Studies in International Economics: The Future of Large, Internationally Active Banks*, Vol. 55, chap. 18, pp. 273–292 (World Scientific).
- Coeurdacier, Nicolas, and Philippe Martin, 2008, “The geography of asset trade and the euro: Insiders and outsiders,” *Journal of the Japanese and International Economies*, Vol. 23, pp. 90–113.
- European Central Bank, 2014, “Financial Integration in Europe,” .
- Fillat, José L., Stefania Garetto, Arthur Smith, and Martin Götz, 2017, “Multinational Banks,” Working Paper.

- Freixas, Xavier, and Jean-Charles Rochet, 2008, *Microeconomics of Banking* (The MIT Press), 2nd ed.
- Giannetti, Mariassunta, and Luc Laeven, 2012, “The flight home effect: Evidence from the syndicated loan market during financial crises,” *Journal of Financial Economics*, Vol. 104, pp. 23–43.
- Gómez-Herrera, Estrella, 2013, “Comparing alternative methods to estimate gravity models of bilateral trade,” *Empirical Economics*, Vol. 44, No. 3, pp. 1087–1111.
- Head, K., and T. Mayer, 2013, “Gravity Equations: Toolkit, Cookbook, Workhorse,” in Gita Gopinath, Elhanan Helpman, and Ken Rogoff (eds.), *Handbook of International Economics*, Vol. 4 (Elsevier).
- Head, K., T. Mayer, and J. Ries, 2010, “The erosion of colonial trade linkages after independence,” *Journal of International Economics*, Vol. 81, No. 1, pp. 1–14.
- Helpman, Elhanan, Marc Melitz, and Yona Rubinstein, 2008, “Estimating Trade Flows: Trading Partners and Trading Volumes,” *The Quarterly Journal of Economics*, Vol. 123, No. 2, pp. 441–487.
- Houston, Joel F., Chen Lin, and Yue Ma, 2012, “Regulatory Arbitrage and International Bank Flows,” *The Journal of Finance*, Vol. 67, No. 5, pp. 1845–1894.
- International Monetary Fund, 2015, “International Banking After the Crisis: Increasingly Local and Safer?” in *Global Financial Stability Report*, chap. 2, pp. 55–91 (International Monetary Fund).
- , 2017, *Financial Integration in Latin America: A Strategy for a New Normal* (International Monetary Fund).
- Ivashina, Victoria, and David Scharfstein, 2010, “Bank lending during the financial crisis of 2008,” *Journal of Financial Economics*, Vol. 97, pp. 319–338.
- Karolyi, G. Andrew, John Sedunov, and Alvaro G. Taboada, 2017, “Cross-Border Bank Flows and Systemic Risk,” Working Paper.
- Karolyi, G. Andrew, and Alvaro G. Taboada, 2015, “Regulatory Arbitrage and Cross-Border Bank Acquisitions,” *The Journal of Finance*, Vol. 70, No. 6, pp. 2395–2450.
- Laeven, Luc, and Fabian Valencia, 2012, “Systemic Banking Crises Database: An Update,” IMF Working Paper 12/163.
- Lane, Philip, and Gian Maria Milesi-Ferretti, 2017, “International financial integration in the aftermath of the global financial crisis,” IMF Working Paper 17/115.
- Martin, Philippe, and Hélène Rey, 2004, “Financial super-markets: size matters for asset trade,” *Journal of International Economics*, Vol. 64, pp. 335–361.

- McCauley, Robert N, Augustin S Benetrix, Patrick McGuire, and Goetz von Peter, 2017, “Financial deglobalisation in banking?” BIS Working Papers No. 650.
- Niepmann, Friederike, 2015, “Banking across borders,” *Journal of International Economics*, Vol. 96, pp. 244–265.
- , 2016, “Banking across borders with heterogeneous banks,” *International Finance Discussion Papers* 1177.
- Okawa, Yohei, and Eric van Wincoop, 2012, “Gravity in International Finance,” *Journal of International Economics*, Vol. 87, pp. 205–215.
- Portes, Richard, and H el ene Rey, 2005, “The determinants of cross-border equity flows,” *Journal of International Economics*, Vol. 65, pp. 269–296.
- Rose, Andrew K., and Tomasz Wieladek, 2014, “Financial protectionism? First evidence,” *The Journal of Finance*, Vol. 69, No. 5, pp. 2127–2149.
- Silva, J.M.C. Santos, and Silvana Tenreyro, 2006, “The Log of Gravity,” *The Review of Economics and Statistics*, Vol. 88, No. 4, pp. 641–658.
- , 2015, “Trading Partners and Trading Flows: Implementing the Helpman-Melitz-Rubinstein Model Empirically,” *Oxford Bulletin of Economics and Statistics*, Vol. 77, No. 1, pp. 93–105.
- Teodoru, Iulia Ruxandra, 2017, “Barriers to Integration in Banking,” in *Financial Integration in Latin America*, chap. 3, pp. 41–58 (International Monetary Fund).

APPENDIX

APPENDIX A. ADDITIONAL ESTIMATION RESULTS

A.1. Alternative Regional Groupings

In addition to UN classification, we use two alternative sets of regional classification to test the robustness of our results. The classifications come from World Bank and IMF, and differ from the baseline classification in only a few countries. World Bank classification is more detailed, separating East Asian countries (Japan, Korea, etc.) from South Asian countries (India, for instance). Turkey belongs to the European group instead of Asia, but is still considered “peripheral” in our lender grouping. Table A1 to Table A8 reports the results, using our baseline estimation methodology. The results are largely robust and in particular support the claim that the role of peripheral lenders as key regional players in the global banking network has evolved and developed significantly since the crisis.

Table A1. WB Regional Grouping: First Stage – Direct Cross-Border

| | (1) No Interactions | (2) Crisis | (3) Peripheral | (4) Peripheral + Crisis |
|-----------------------------------|------------------------|-------------------------|-----------------------|----------------------------|
| Region | 0.0333*** (0.0128) | 0.106*** (0.0151) | -0.0329** (0.0130) | 0.0588*** (0.0164) |
| Region * Crisis | | -0.0708*** (0.00906) | | -0.0794*** (0.00969) |
| Region * Post-crisis | | -0.0866*** (0.00858) | | -0.115*** (0.00889) |
| Region * Peripheral | | | 0.192*** (0.0207) | 0.0714*** (0.0237) |
| Region * Peripheral * Crisis | | | | 0.0650*** (0.0251) |
| Region * Peripheral * Post-crisis | | | | 0.171*** (0.0236) |
| N | 139216 | 139216 | 139216 | 139216 |
| Pseudo R-sq | 0.224 | 0.226 | 0.228 | 0.231 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Table A2. WB Regional Grouping: First Stage – Local Affiliate

| | (1) No Interactions | (2) Crisis | (3) Peripheral | (4) Peripheral + Crisis |
|-----------------------------------|------------------------|-----------------------|----------------------|----------------------------|
| Region | 0.0480*** (0.0160) | 0.0602*** (0.0179) | 0.0433** (0.0182) | 0.0597*** (0.0201) |
| Region * Crisis | | -0.00573 (0.00735) | | -0.00236 (0.00776) |
| Region * Post-crisis | | -0.0161* (0.00835) | | -0.0236*** (0.00863) |
| Region * Peripheral | | | 0.0117 (0.0199) | -0.0126 (0.0242) |
| Region * Peripheral * Crisis | | | | -0.0295* (0.0166) |
| Region * Peripheral * Post-crisis | | | | 0.0542** (0.0255) |
| N | 91996 | 91996 | 91996 | 91996 |
| Pseudo R-sq | 0.262 | 0.262 | 0.262 | 0.263 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Table A3. IMF Regional Grouping: First Stage – Direct Cross-Border

| | (1) No Interactions | (2) Crisis | (3) Peripheral | (4) Peripheral + Crisis |
|-----------------------------------|------------------------|-------------------------|----------------------|----------------------------|
| Region | 0.0462*** (0.0111) | 0.0986*** (0.0130) | -0.00328 (0.0117) | 0.0748*** (0.0148) |
| Region * Crisis | | -0.0474*** (0.00945) | | -0.0588*** (0.0102) |
| Region * Post-crisis | | -0.0633*** (0.00872) | | -0.0980*** (0.00934) |
| Region * Peripheral | | | 0.159*** (0.0183) | 0.0537*** (0.0207) |
| Region * Peripheral * Crisis | | | | 0.0562** (0.0230) |
| Region * Peripheral * Post-crisis | | | | 0.152*** (0.0211) |
| N | 139216 | 139216 | 139216 | 139216 |
| Pseudo R-sq | 0.225 | 0.226 | 0.228 | 0.230 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Table A4. IMF Regional Grouping: First Stage – Local Affiliate

| | (1) No Interactions | (2) Crisis | (3) Peripheral | (4) Peripheral + Crisis |
|-----------------------------------|------------------------|-----------------------|-----------------------|----------------------------|
| Region | 0.0449*** (0.0114) | 0.0553*** (0.0133) | 0.0269** (0.0127) | 0.0442*** (0.0148) |
| Region * Crisis | | -0.00333 (0.00739) | | -0.000983 (0.00781) |
| Region * Post-crisis | | -0.0145* (0.00823) | | -0.0260*** (0.00859) |
| Region * Peripheral | | | 0.0543*** (0.0206) | 0.0188 (0.0232) |
| Region * Peripheral * Crisis | | | | -0.0143 (0.0159) |
| Region * Peripheral * Post-crisis | | | | 0.0620*** (0.0217) |
| N | 91996 | 91996 | 91996 | 91996 |
| Pseudo R-sq | 0.262 | 0.263 | 0.264 | 0.265 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Table A5. WB Regional Grouping: Second Stage – Direct Cross-Border

| | (1) No Interactions | (2) Crisis | (3) Peripheral | (4) Peripheral + Crisis |
|-----------------------------------|------------------------|----------------------|---------------------|----------------------------|
| Region | 0.479*** (0.108) | 0.451*** (0.123) | 0.114 (0.124) | 0.208 (0.127) |
| Region * Crisis | | -0.00625 (0.0676) | | -0.0754 (0.0735) |
| Region * Post-crisis | | 0.0291 (0.0781) | | -0.201** (0.0998) |
| Region * Peripheral | | | 0.839*** (0.207) | 0.506*** (0.170) |
| Region * Peripheral * Crisis | | | | 0.246* (0.136) |
| Region * Peripheral * Post-crisis | | | | 0.690*** (0.165) |
| N | 37625 | 37625 | 37625 | 37625 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Table A6. WB Regional Grouping: Second Stage – Local Affiliates

| | (1) No Interactions | (2) Crisis | (3) Peripheral | (4) Peripheral + Crisis |
|-----------------------------------|------------------------|-------------------|---------------------|----------------------------|
| Region | 0.306 (0.388) | 0.271 (0.441) | -0.138 (0.441) | -0.123 (0.526) |
| Region * Crisis | | 0.127 (0.151) | | 0.161 (0.153) |
| Region * Post-crisis | | 0.0461 (0.204) | | -0.0243 (0.257) |
| Region * Peripheral | | | 1.190*** (0.425) | 1.348*** (0.520) |
| Region * Peripheral * Crisis | | | | -0.541 (0.448) |
| Region * Peripheral * Post-crisis | | | | -0.0845 (0.580) |
| N | 11044 | 11044 | 11044 | 11044 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Table A7. IMF Regional Grouping: Second Stage – Direct Cross-Border

| | (1) No Interactions | (2) Crisis | (3) Peripheral | (4) Peripheral + Crisis |
|-----------------------------------|------------------------|---------------------|---------------------|----------------------------|
| Region | 0.250*** (0.0789) | 0.232** (0.0933) | 0.0917 (0.0858) | 0.191* (0.101) |
| Region * Crisis | | 0.0178 (0.0623) | | -0.0416 (0.0683) |
| Region * Post-crisis | | 0.00540 (0.0710) | | -0.198** (0.0927) |
| Region * Peripheral | | | 0.494*** (0.166) | 0.158 (0.145) |
| Region * Peripheral * Crisis | | | | 0.239** (0.121) |
| Region * Peripheral * Post-crisis | | | | 0.602*** (0.148) |
| N | 37625 | 37625 | 37625 | 37625 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Table A8. IMF Regional Grouping: Second Stage – Local Affiliates

| | (1) No Interactions | (2) Crisis | (3) Peripheral | (4) Peripheral + Crisis |
|-----------------------------------|------------------------|-------------------|-------------------|----------------------------|
| Region | -0.153 (0.307) | -0.309 (0.361) | -0.325 (0.328) | -0.291 (0.404) |
| Region * Crisis | | 0.177 (0.150) | | 0.200 (0.157) |
| Region * Post-crisis | | 0.194 (0.200) | | -0.0116 (0.267) |
| Region * Peripheral | | | 0.522 (0.429) | 0.360 (0.454) |
| Region * Peripheral * Crisis | | | | -0.272 (0.309) |
| Region * Peripheral * Post-crisis | | | | 0.527 (0.538) |
| N | 11044 | 11044 | 11044 | 11044 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Notes: Table A1 to Table A8 report the robustness check of our estimations, using alternative classification to define the common region indicator. Only the coefficients related to the common region variable are shown. World Bank and IMF classifications are used in lieu of the baseline UN classification (see Table 2). Standard errors are clustered at country-pair level. For the first stage, all coefficients are marginal effects. Second-stage estimations are carried out using maximum likelihood.

A.2. Alternative Methods

We provide four alternative methods to estimate the gravity equation. First, we use non-linear least squares, as originally proposed in [Helpman, Melitz, and Rubinstein \(2008\)](#), to estimate the second stage equation ((20)). Our MLE procedure obtains point estimates very close to those shown in Table A9 and A10, while being more precisely estimated. We calculate the R-square measure of our second-stage non-linear least squares model using the squared correlation coefficient between predicted log flows and the actual flows. In an OLS setting, the squared correlation equals the traditionally defined R-square. The R-square of our model outperforms similar attempts in the literature, such as [Houston, Lin, and Ma \(2012\)](#), and is close to the results obtained from trade gravity equation ([Brei and von Peter, 2017](#); [Helpman, Melitz, and Rubinstein, 2008](#)).

Next, as in [Helpman, Melitz, and Rubinstein \(2008\)](#), we relax the assumption that latent inverse bank productivity follows a Pareto distribution. Instead, the distribution is approximated by polynomials of $z_{ij} = \hat{z}_{ij}^{x^*} + \widehat{\eta}_{ij}^{x^*}$ obtained from the Probit first stage. We implement this semi-parametric procedure using OLS, and the results stabilize at quintics for direct cross-border and at quartics for local affiliate flows. Table A11 reports the second-stage results. Using UN regional definition, this method yields results that are also similar to the baseline.

Other methods that properly take into account the zeroes and non-connections in the gravity estimation have been proposed in the literature. Among others, Poisson pseudo-maximum likelihood estimator (PPML), as in [Silva and Tenreyro \(2006\)](#), can be used to estimate Equation (5) and (6) directly. To see the rationale, rewrite Equation (5) or (6) as

$$L_{ij}^X = A\Lambda_i^\beta \Phi_j^\theta \tau_{ij}^\eta U_{ij} \quad (23)$$

where $X \in \{CB, LC\}$ and U_{ij} is the error term. This method enables us to estimate the parameters $\{\beta, \theta, \eta\}$ using L_{ij}^X in its absolute level and thus preserve zeroes in the data. This method is fast to estimate, but does not have a model-based foundation to account for the unobserved productivity component responsible for the zeroes. Table A12 reports the results. While the results are not directly comparable to our single-stage baseline estimates, the PPML estimates for direct cross-border flows are largely in line with the baseline results. The coefficient of the triple interaction between common region, peripheral lender and post-crisis is large, positive and strongly significant, consistent with the first-stage estimate in Table 3. For local affiliate flows that come with higher barriers to entry, however, the PPML estimators are not significant except for legal system, borrower log GDP and log geographical distance, possibly due

Table A9. Non-linear Least Squares Regression: Second Stage – Direct Cross Border

| | (1) No Interactions | (2) Crisis | (3) Peripheral | (4) Peripheral + Crisis |
|-----------------------------------|------------------------|----------------------|----------------------|----------------------------|
| Region | -0.0591 (0.0846) | -0.0479 (0.129) | -0.405*** (0.157) | -0.318*** (0.110) |
| Region * Crisis | | 0.0102 (0.121) | | -0.0138 (0.154) |
| Region * Post-crisis | | -0.0296 (0.150) | | -0.153 (0.218) |
| Region * Peripheral | | | 0.806*** (0.301) | 0.606*** (0.190) |
| Region * Peripheral * Crisis | | | | 0.102 (0.178) |
| Region * Peripheral * Post-Crisis | | | | 0.315 (0.251) |
| Language | 0.157 (0.106) | 0.152 (0.106) | 0.190* (0.107) | 0.188* (0.108) |
| Colonizer | 0.464*** (0.159) | 0.453*** (0.159) | 0.392** (0.156) | 0.387** (0.157) |
| Legal | 0.153** (0.0684) | 0.150** (0.0690) | 0.150** (0.0667) | 0.148** (0.0673) |
| Colonial Relation | 0.840*** (0.192) | 0.824*** (0.193) | 0.814*** (0.175) | 0.802*** (0.174) |
| FTA | 0.131 (0.0811) | 0.130 (0.0815) | 0.215** (0.0934) | 0.210** (0.0926) |
| Log Distance | -0.984*** (0.127) | -0.966*** (0.128) | -0.964*** (0.112) | -0.940*** (0.116) |
| Lender Log GDP | 0.809** (0.396) | 0.809** (0.400) | 0.790** (0.378) | 0.758** (0.329) |
| Borrower Log GDP | 0.693*** (0.242) | 0.674*** (0.225) | 0.683*** (0.234) | 0.653*** (0.190) |
| δ | 0.148 (0.725) | 0.157 (0.752) | 0.149 (0.713) | 0.211 (0.804) |
| β | 0.803* (0.485) | 0.717 (0.490) | 0.819* (0.452) | 0.763* (0.444) |
| N | 37625 | 37625 | 37625 | 37625 |
| R-squared | 0.678 | 0.678 | 0.680 | 0.680 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Notes: This table reports the second-stage non-linear least squares regression result, with log cross-border flow as dependent variable. Standard errors are clustered at country-pair level. δ and β refer to, respectively, the parameter of the non-linear adjustment term and the estimated inverse Mill's ratio term in Equation (20).

to PPML's inability to dissect the overall effect of regional and gravitational factors into link formation and link intensity components.

Table A10. Non-linear Least Squares Regression: Second Stage – Local Affiliate

| | (1) No Interactions | (2) Crisis | (3) Peripheral | (4) Peripheral + Crisis |
|-----------------------------------|------------------------|----------------------|----------------------|----------------------------|
| Region | -0.0918 (0.290) | -0.269 (0.317) | -0.318 (0.415) | -0.414 (0.405) |
| Region * Crisis | | 0.235 (0.152) | | 0.208 (0.156) |
| Region * Post-crisis | | 0.215 (0.203) | | -0.0713 (0.257) |
| Region * Peripheral | | | 0.661 (0.689) | 0.0585 (0.725) |
| Region * Peripheral * Crisis | | | | 0.471 (0.401) |
| Region * Peripheral * Post-Crisis | | | | 1.367** (0.568) |
| Language | 0.255 (0.267) | 0.254 (0.268) | 0.283 (0.270) | 0.341 (0.273) |
| Colonizer | 0.673 (0.603) | 0.658 (0.605) | 0.384 (0.644) | 0.379 (0.632) |
| Legal | 0.360* (0.207) | 0.355* (0.213) | 0.387* (0.207) | 0.477** (0.225) |
| Colonial Relation | 1.516*** (0.405) | 1.501*** (0.418) | 1.513*** (0.389) | 1.739*** (0.442) |
| FTA | 0.0889 (0.256) | 0.0836 (0.258) | 0.155 (0.275) | 0.291 (0.292) |
| Log Distance | -1.363*** (0.287) | -1.352*** (0.297) | -1.397*** (0.288) | -1.539*** (0.319) |
| Lender Log GDP | -0.271 (0.306) | -0.284 (0.314) | -0.227 (0.306) | -0.178 (0.318) |
| Borrower Log GDP | 1.322*** (0.330) | 1.327*** (0.325) | 1.304*** (0.329) | 1.419*** (0.327) |
| δ | 5.253*** (1.429) | 5.335*** (1.466) | 5.212*** (1.433) | 4.329*** (1.590) |
| β | 1.387* (0.751) | 1.350* (0.795) | 1.454** (0.736) | 1.982** (0.917) |
| N | 11044 | 11044 | 11044 | 11044 |
| R-squared | 0.546 | 0.546 | 0.547 | 0.548 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Notes: This table reports the second-stage non-linear least squares regression result on the baseline model, with log local affiliate flow as dependent variable. Standard errors are clustered at country-pair level. δ and β refer to, respectively, the parameter of the non-linear adjustment term and the estimated inverse Mill's ratio term in Equation (20).

Finally, we estimate the log-linear counterpart of Equation (23) using OLS, and report the output in Table A13. Overall, the results are similar to our baseline non-linear procedures. Without accounting for the positive effect of selection of bank lending targets and the propor-

tion of banks with international subsidiaries on local affiliate flows, however, OLS estimates are biased.

Table A11. Semiparametric Regression: Second Stage

| | Direct Cross-Border | | | | Local Affiliate | | | |
|-----------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Region | -0.0740 (0.0840) | 0.0204 (0.141) | -0.412** (0.167) | -0.270** (0.111) | -0.0986 (0.326) | -0.279 (0.388) | -0.353 (0.441) | -0.415 (0.411) |
| Region * Crisis | | -0.0817 (0.137) | | -0.0661 (0.166) | | 0.233 (0.152) | | 0.210 (0.156) |
| Region * Post-crisis | | -0.161 (0.177) | | -0.252 (0.241) | | 0.212 (0.240) | | 0.0113 (0.296) |
| Region * Peripheral | | | 0.815** (0.320) | 0.504*** (0.192) | | | 0.751 (0.986) | -0.184 (0.831) |
| Region * Peripheral * Crisis | | | | 0.176 (0.189) | | | | 0.478 (0.401) |
| Region * Peripheral * Post-Crisis | | | | 0.468* (0.278) | | | | 1.242** (0.618) |
| Language | 0.182* (0.106) | 0.180* (0.106) | 0.202* (0.108) | 0.198* (0.107) | 0.268 (0.277) | 0.263 (0.277) | 0.307 (0.284) | 0.304 (0.282) |
| Colonizer | 0.515*** (0.160) | 0.511*** (0.161) | 0.418*** (0.157) | 0.419*** (0.157) | 0.664 (0.640) | 0.639 (0.642) | 0.363 (0.625) | 0.341 (0.618) |
| Legal | 0.173** (0.0706) | 0.173** (0.0714) | 0.150** (0.0686) | 0.153** (0.0682) | 0.392 (0.309) | 0.379 (0.310) | 0.420 (0.289) | 0.402 (0.281) |
| Colonial Relation | 0.962*** (0.200) | 0.952*** (0.203) | 0.848*** (0.182) | 0.837*** (0.179) | 1.573** (0.723) | 1.535** (0.723) | 1.580** (0.631) | 1.532** (0.605) |
| FTA | 0.119 (0.0798) | 0.119 (0.0801) | 0.210** (0.0942) | 0.197** (0.0928) | 0.112 (0.338) | 0.0967 (0.337) | 0.201 (0.373) | 0.189 (0.359) |
| Log Distance | -1.060*** (0.141) | -1.065*** (0.145) | -0.972*** (0.124) | -0.980*** (0.125) | -1.386*** (0.434) | -1.363*** (0.437) | -1.423*** (0.401) | -1.404*** (0.387) |
| Lender Log GDP | 0.975** (0.432) | 0.977** (0.438) | 0.809** (0.404) | 0.625* (0.339) | -0.242 (1.125) | -0.300 (1.097) | -0.153 (1.031) | -0.646 (0.842) |
| Borrower Log GDP | 0.782*** (0.261) | 0.745*** (0.241) | 0.680*** (0.247) | 0.567*** (0.193) | 1.343* (0.768) | 1.324* (0.715) | 1.342* (0.703) | 1.128** (0.561) |
| z | -77.87*** (15.36) | -77.49*** (13.65) | -73.69*** (15.45) | -70.08*** (13.52) | -77.24 (49.45) | -74.25 (48.80) | -70.02 (46.19) | -68.40 (43.58) |
| z^2 | 137.3*** (27.85) | 135.4*** (23.49) | 132.8*** (28.32) | 124.2*** (23.85) | 135.9* (77.73) | 131.0* (75.77) | 119.6 (74.44) | 119.1* (69.36) |
| z^3 | -115.5*** (25.65) | -112.5*** (20.85) | -113.1*** (26.11) | -103.7*** (21.30) | -98.14* (59.58) | -94.08 (57.40) | -83.34 (58.01) | -84.28 (53.14) |
| z^4 | 46.32*** (11.38) | 44.48*** (8.936) | 45.90*** (11.57) | 41.11*** (9.137) | 25.21 (17.19) | 23.98 (16.37) | 20.52 (16.95) | 21.09 (15.22) |
| z^5 | -7.163*** (1.934) | -6.770*** (1.467) | -7.163*** (1.960) | -6.270*** (1.498) | | | | |
| N | 37625 | 37625 | 37625 | 37625 | 11044 | 11044 | 11044 | 11044 |
| R-sq | 0.680 | 0.681 | 0.682 | 0.682 | 0.548 | 0.548 | 0.548 | 0.550 |
| Lender + Borrower FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES | YES | YES | YES | YES |

Notes: This table reports the second-stage regression results, using polynomials of $z = \hat{z}_{ij}^* + \hat{\eta}_{ij}^*$, obtained from first-stage Probit to approximate arbitrary distribution of latent bank productivity. Estimations are conducted using OLS. For cross-border, estimates stabilize at fifth power, while for local affiliate flows, estimates stabilize at quartics. Standard errors are clustered at country-pair level.

Table A12. Poisson Regression

| | Direct Cross-Border | | | | Local Affiliate | | | |
|-----------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|----------------------|----------------------|----------------------|----------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Region | -0.206 (0.147) | 0.253 (0.184) | -0.493*** (0.147) | 0.0156 (0.187) | -0.0968 (0.384) | -0.206 (0.413) | -0.317 (0.372) | -0.418 (0.392) |
| Region * Crisis | | -0.545*** (0.129) | | -0.538*** (0.129) | | 0.193 (0.311) | | 0.194 (0.315) |
| Region * Post-crisis | | -0.783*** (0.187) | | -0.917*** (0.194) | | 0.151 (0.487) | | 0.130 (0.494) |
| Region * Peripheral | | | 1.301*** (0.253) | 0.523** (0.252) | | | 1.439 (1.076) | 1.002 (1.288) |
| Region * Peripheral * Crisis | | | | 0.144 (0.464) | | | | 0.367 (0.401) |
| Region * Peripheral * Post-Crisis | | | | 1.448*** (0.338) | | | | 0.756 (0.650) |
| Language | 0.106 (0.129) | 0.112 (0.130) | 0.0594 (0.118) | 0.0583 (0.117) | -0.280 (0.447) | -0.281 (0.448) | -0.330 (0.439) | -0.332 (0.439) |
| Colonizer | 0.886** (0.365) | 0.914** (0.368) | 0.265 (0.344) | 0.277 (0.342) | -0.421 (0.782) | -0.421 (0.780) | -1.328 (1.017) | -1.330 (0.996) |
| Legal | 0.210** (0.0942) | 0.207** (0.0944) | 0.231** (0.0900) | 0.231*** (0.0897) | 0.907*** (0.238) | 0.907*** (0.238) | 0.914*** (0.234) | 0.914*** (0.234) |
| Colonial Relation | 0.176 (0.159) | 0.189 (0.159) | 0.164 (0.153) | 0.172 (0.152) | 0.0757 (0.441) | 0.0745 (0.443) | 0.0918 (0.444) | 0.0903 (0.446) |
| FTA | 0.142 (0.142) | 0.149 (0.142) | 0.165 (0.136) | 0.164 (0.135) | -0.212 (0.450) | -0.207 (0.439) | -0.184 (0.436) | -0.179 (0.425) |
| Log Distance | -0.372*** (0.0930) | -0.361*** (0.0931) | -0.405*** (0.0909) | -0.397*** (0.0906) | -0.794*** (0.209) | -0.797*** (0.209) | -0.808*** (0.209) | -0.811*** (0.209) |
| Lender Log GDP | 0.184 (0.462) | 0.188 (0.488) | 0.178 (0.458) | -0.153 (0.496) | -0.670 (0.787) | -0.696 (0.805) | -0.688 (0.789) | -0.870 (0.835) |
| Borrower Log GDP | 1.175*** (0.249) | 1.008*** (0.277) | 1.175*** (0.249) | 0.894*** (0.271) | 1.804** (0.808) | 1.841** (0.773) | 1.806** (0.808) | 1.817** (0.770) |
| N | 190936 | 190936 | 190936 | 190936 | 127700 | 127700 | 127700 | 127700 |
| R-sq | 0.325 | 0.337 | 0.324 | 0.338 | 0.184 | 0.189 | 0.184 | 0.188 |
| Lender + Borrower FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES | YES | YES | YES | YES |

Notes: This table reports the output of Poisson pseudo-maximum likelihood (PPML) estimates. Dependent variable is now the absolute level of cross-border / local affiliate flows, which include zeroes (non-connection) as well. Standard errors are clustered at country-pair level.

Table A13. OLS Regression

| | Direct Cross-Border | | | | Local Affiliate | | | |
|-----------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|----------------------|----------------------|----------------------|----------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Region | -0.0251 (0.0784) | 0.0491 (0.0838) | -0.278*** (0.0960) | -0.0906 (0.104) | -0.343 (0.285) | -0.458 (0.299) | -0.660* (0.386) | -0.645 (0.394) |
| Region * Crisis | | -0.0394 (0.0572) | | -0.100 (0.0637) | | 0.252* (0.143) | | 0.243 (0.150) |
| Region * Post-crisis | | -0.124* (0.0632) | | -0.319*** (0.0736) | | 0.115 (0.165) | | -0.133 (0.174) |
| Region * Peripheral | | | 0.535*** (0.140) | 0.193 (0.152) | | | 0.763 (0.491) | -0.0217 (0.561) |
| Region * Peripheral * Crisis | | | | 0.220** (0.105) | | | | 0.178 (0.302) |
| Region * Peripheral * Post-Crisis | | | | 0.545*** (0.115) | | | | 1.182*** (0.364) |
| Language | 0.110 (0.0919) | 0.109 (0.0919) | 0.120 (0.0911) | 0.121 (0.0911) | 0.351 (0.262) | 0.352 (0.262) | 0.366 (0.262) | 0.364 (0.262) |
| Colonizer | 0.428*** (0.136) | 0.424*** (0.136) | 0.404*** (0.137) | 0.405*** (0.137) | 0.928 (0.573) | 0.917 (0.573) | 0.739 (0.601) | 0.742 (0.594) |
| Legal | 0.0997* (0.0547) | 0.101* (0.0547) | 0.0954* (0.0545) | 0.0970* (0.0545) | 0.291* (0.177) | 0.294* (0.177) | 0.312* (0.177) | 0.312* (0.177) |
| Colonial Relation | 0.886*** (0.119) | 0.887*** (0.119) | 0.874*** (0.118) | 0.873*** (0.118) | 1.406*** (0.310) | 1.403*** (0.309) | 1.388*** (0.306) | 1.387*** (0.305) |
| FTA | 0.129* (0.0771) | 0.127 (0.0771) | 0.201** (0.0790) | 0.194** (0.0792) | 0.189 (0.207) | 0.188 (0.207) | 0.291 (0.222) | 0.299 (0.223) |
| Log Distance | -0.975*** (0.0569) | -0.976*** (0.0570) | -0.959*** (0.0572) | -0.957*** (0.0572) | -1.569*** (0.179) | -1.572*** (0.179) | -1.575*** (0.182) | -1.580*** (0.182) |
| Lender Log GDP | 0.939*** (0.124) | 0.948*** (0.125) | 0.957*** (0.124) | 0.773*** (0.131) | 0.288 (0.473) | 0.279 (0.473) | 0.318 (0.470) | -0.149 (0.471) |
| Borrower Log GDP | 0.601*** (0.0919) | 0.576*** (0.0931) | 0.612*** (0.0917) | 0.494*** (0.0941) | 1.573*** (0.279) | 1.562*** (0.282) | 1.545*** (0.281) | 1.305*** (0.288) |
| N | 45183 | 45183 | 45183 | 45183 | 12691 | 12691 | 12691 | 12691 |
| R-sq | 0.664 | 0.664 | 0.665 | 0.665 | 0.536 | 0.536 | 0.537 | 0.538 |
| Lender + Borrower FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES | YES | YES | YES | YES |

Notes: This table reports the OLS estimation of our gravity model specification (i.e. without including non-linear adjustment and Heckman adjustment term). Standard errors are clustered at country-pair level.

A.3. Gravity Factor Interactions, Alternative Definition of Flow

A.3.1. Interactions between Gravity Factors and Regional Proximity: Pre-crisis and Crisis Sample

In section 6 of the main text, we use the interactions between the regional dummy and other static gravity factors in our model to show that regional proximity serves as a unique determinant of global banking flows. Table [A14](#) and [A15](#) repeat the same exercise on the sample spanning the period 2006Q1–2009Q2.

A.3.2. Alternative Definition of Flow

In the main text we define flow as the positive year-over-year change in stock. Past studies, such as [Houston, Lin, and Ma \(2012\)](#), also use growth rate, often expressed in log difference, as flow variable. As this calculation introduces negative flow, not adaptable to our empirical framework, we use OLS to estimate the same gravity equation, except that in this case, neither a non-linear term nor a Heckman selection term is present. Table [A16](#) reports the result. While the common region dummy and its interactions are still highly significant, supporting our findings in terms of regionalization, other gravity variables have less explanatory power, or come with counterintuitive sign. In particular, geographical distance is insignificant. The model fit is also poor in general, with the R-squared measure less than 0.1 in all specifications. Using growth rate as the flow variable, as a result, may not be fully compatible with time-invariant factors in the traditional gravity framework.

Table A14. Region as an Interaction Term: First Stage, 2006Q1–2009Q2 Sample

| | Direct Cross-Border | | Local Affiliate | |
|----------------------------|------------------------|------------------------|------------------------|------------------------|
| | (1) | (2) | (3) | (4) |
| Region | -0.316*** (0.00801) | -0.312*** (0.0114) | -0.0915 (0.142) | -0.0892 (0.141) |
| Region * Peripheral | | 0.120*** (0.0277) | | 0.108*** (0.0448) |
| Language | 0.0267* (0.0143) | 0.0308** (0.0143) | 0.0236 (0.0216) | 0.0308 (0.0221) |
| Colonizer | 0.0208 (0.0294) | 0.0394 (0.0304) | 0.0401 (0.0653) | 0.0674 (0.0713) |
| Legal | 0.0247*** (0.00897) | 0.0260*** (0.00900) | 0.0277** (0.0125) | 0.0289** (0.0126) |
| Colonial Relation | 0.0852*** (0.0196) | 0.0813*** (0.0195) | 0.0592** (0.0282) | 0.0537* (0.0279) |
| FTA | 0.0351** (0.0168) | 0.0404** (0.0170) | 0.0292 (0.0220) | 0.0376 (0.0230) |
| Log Distance | -0.120*** (0.0112) | -0.108*** (0.0113) | -0.0698*** (0.0155) | -0.0684*** (0.0157) |
| Lender Log GDP | 0.193*** (0.0480) | 0.192*** (0.0482) | -0.0525 (0.0458) | -0.0544 (0.0461) |
| Borrower Log GDP | 0.214*** (0.0295) | 0.214*** (0.0295) | 0.175*** (0.0331) | 0.176*** (0.0330) |
| Language * Region | 0.0161 (0.0273) | 0.000367 (0.0265) | -0.0811*** (0.0215) | -0.0885*** (0.0198) |
| Colonizer * Region | -0.00538 (0.0394) | -0.0530 (0.0373) | 0.0180 (0.0707) | -0.0361 (0.0583) |
| Legal * Region | 0.0118 (0.0169) | 0.00631 (0.0166) | 0.00781 (0.0208) | 0.00346 (0.0203) |
| Colonial Relation * Region | -0.108*** (0.0350) | -0.105*** (0.0348) | -0.0216 (0.0534) | -0.0165 (0.0542) |
| FTA * Region | -0.0618*** (0.0220) | -0.0402* (0.0234) | -0.0536** (0.0268) | -0.0494* (0.0266) |
| Log Distance * Region | 0.0881*** (0.0161) | 0.0742*** (0.0162) | 0.0160 (0.0206) | 0.0117 (0.0203) |
| Concentration | 0.0352 (0.0503) | 0.0341 (0.0501) | 0.0704 (0.0703) | 0.0760 (0.0703) |
| Overhead | -0.0281 (0.0192) | -0.0339* (0.0193) | -0.0677 (0.0539) | -0.0726 (0.0537) |
| Entry Cost | | | -0.0224 (0.0254) | -0.0277 (0.0261) |
| Entry Time | | | -0.0173 (0.0154) | -0.0211 (0.0157) |
| N | 55854 | 55854 | 32500 | 32500 |
| Pseudo R-sq | 0.308 | 0.309 | 0.320 | 0.321 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Notes: This table reports the first-stage Probit regression result on the baseline model, with the common region indicator serving as a gravity factor. Its interaction with traditional, static gravity factors are also included. Standard errors are clustered at country-pair level. All coefficients are marginal effects.

Table A15. Region as an Interaction Term: Second Stage, 2006Q1–2009Q2 Sample

| | Direct Cross-Border | | Local Affiliate | |
|----------------------------|----------------------|----------------------|----------------------|----------------------|
| | (1) | (2) | (3) | (4) |
| Region | -0.945 (1.277) | -0.509 (1.253) | 5.307* (3.206) | 5.505* (3.106) |
| Region * Peripheral | | 1.303*** (0.206) | | 1.039 (0.924) |
| Language | 0.112 (0.124) | 0.197 (0.123) | 0.454 (0.374) | 0.483 (0.396) |
| Colonizer | 0.479** (0.225) | 0.726*** (0.225) | 2.406*** (0.900) | 2.635*** (0.927) |
| Legal | 0.0882 (0.0741) | 0.0927 (0.0744) | 0.322 (0.348) | 0.285 (0.361) |
| Colonial Relation | 1.100*** (0.136) | 1.045*** (0.135) | 1.558*** (0.510) | 1.388*** (0.509) |
| FTA | 0.330*** (0.118) | 0.399*** (0.118) | 0.491 (0.396) | 0.529 (0.434) |
| Log Distance | -1.090*** (0.130) | -0.984*** (0.129) | -1.352** (0.587) | -1.186** (0.588) |
| Lender Log GDP | 0.748*** (0.190) | 0.610*** (0.195) | -0.737 (0.786) | -0.561 (0.831) |
| Borrower Log GDP | 0.716*** (0.195) | 0.702*** (0.202) | 2.091 (1.300) | 1.708 (1.349) |
| Language * Region | -0.0720 (0.214) | -0.260 (0.205) | -1.185 (0.912) | -1.088 (0.998) |
| Colonizer * Region | -0.333 (0.291) | -0.967*** (0.308) | -2.374** (1.105) | -3.133*** (1.148) |
| Legal * Region | 0.241* (0.134) | 0.200 (0.133) | 0.507 (0.403) | 0.464 (0.406) |
| Colonial Relation * Region | -0.620 (0.415) | -0.495 (0.409) | 0.306 (0.638) | 0.393 (0.633) |
| FTA * Region | -0.487** (0.209) | -0.304 (0.201) | -0.714 (0.820) | -0.555 (0.847) |
| Log Distance * Region | 0.125 (0.149) | 0.0120 (0.145) | -0.625* (0.373) | -0.698** (0.355) |
| δ | 0.212*** (0.0524) | 0.282* (0.159) | 1.504 (1.866) | 1.975 (1.893) |
| β | 1.409*** (0.252) | 1.471*** (0.252) | 1.650 (1.293) | 1.268 (1.316) |
| σ | 1.428*** (0.0148) | 1.422*** (0.0148) | 2.047*** (0.0474) | 2.044*** (0.0475) |
| N | 17404 | 17404 | 5070 | 5070 |
| Lender + Borrower FE | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES |

Notes: This table reports the second-stage maximum likelihood regression result on the baseline model, with the common region indicator serving as a gravity factor. Its interaction with traditional, static gravity factors are also included. Standard errors are clustered at country-pair level. δ and β refer to, respectively, the parameter of the non-linear adjustment term and the estimated inverse Mill's ratio term in Equation (20). σ is the estimated standard deviation of the disturbance term, e_{ij} , in Equation (20).

Table A16. OLS Regression: Log Difference

| | Direct Cross-Border | | | | Local Affiliate | | | |
|-----------------------------------|-----------------------|------------------------|-----------------------|-----------------------|----------------------|-----------------------|----------------------|-----------------------|
| | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) |
| Region | 0.0310** (0.0136) | 0.0990*** (0.0240) | 0.0289* (0.0166) | 0.162*** (0.0277) | 0.0498 (0.0419) | 0.213*** (0.0728) | 0.0304 (0.0597) | 0.256*** (0.0883) |
| Region * Crisis | | -0.0696** (0.0308) | | -0.110*** (0.0320) | | -0.0816 (0.0908) | | -0.124 (0.0961) |
| Region * Post-crisis | | -0.0858*** (0.0258) | | -0.174*** (0.0287) | | -0.223*** (0.0769) | | -0.322*** (0.0855) |
| Region * Peripheral | | | 0.00450 (0.0230) | -0.207*** (0.0451) | | | 0.0460 (0.0843) | -0.359*** (0.126) |
| Region * Peripheral * Crisis | | | | 0.137** (0.0606) | | | | 0.289* (0.162) |
| Region * Peripheral * Post-Crisis | | | | 0.281*** (0.0497) | | | | 0.539*** (0.130) |
| Language | -0.00382 (0.0136) | -0.00380 (0.0136) | -0.00372 (0.0136) | -0.00341 (0.0135) | 0.0447 (0.0438) | 0.0475 (0.0439) | 0.0454 (0.0438) | 0.0441 (0.0442) |
| Colonizer | 0.00464 (0.0260) | 0.00394 (0.0260) | 0.00443 (0.0261) | 0.00584 (0.0261) | -0.308*** (0.108) | -0.297*** (0.107) | -0.320*** (0.108) | -0.323*** (0.108) |
| Legal | 0.0119 (0.00844) | 0.0120 (0.00844) | 0.0119 (0.00845) | 0.0118 (0.00845) | 0.0118 (0.0288) | 0.00909 (0.0288) | 0.0132 (0.0292) | 0.0131 (0.0293) |
| Colonial Relation | -0.00628 (0.0148) | -0.00628 (0.0148) | -0.00640 (0.0148) | -0.00679 (0.0148) | -0.0483 (0.0543) | -0.0500 (0.0535) | -0.0496 (0.0544) | -0.0488 (0.0534) |
| FTA | 0.00534 (0.0137) | 0.00453 (0.0137) | 0.00590 (0.0140) | 0.00234 (0.0140) | -0.0322 (0.0402) | -0.0348 (0.0401) | -0.0255 (0.0430) | -0.0216 (0.0431) |
| Log Distance | -0.00623 (0.00832) | -0.00619 (0.00832) | -0.00613 (0.00833) | -0.00699 (0.00832) | -0.0421 (0.0263) | -0.0412 (0.0263) | -0.0426 (0.0264) | -0.0429 (0.0264) |
| Lender Log GDP | 0.0837** (0.0405) | 0.0865** (0.0406) | 0.0839** (0.0405) | 0.00611 (0.0432) | 0.383*** (0.139) | 0.358** (0.139) | 0.383*** (0.139) | 0.196 (0.150) |
| Borrower Log GDP | 0.0662** (0.0288) | 0.0535* (0.0295) | 0.0663** (0.0288) | 0.0162 (0.0305) | 0.247** (0.104) | 0.184* (0.106) | 0.245** (0.104) | 0.108 (0.110) |
| N | 102641 | 102641 | 102641 | 102641 | 20164 | 20164 | 20164 | 20164 |
| R-sq | 0.038 | 0.038 | 0.038 | 0.039 | 0.064 | 0.065 | 0.064 | 0.066 |
| Lender + Borrower FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Quarter FE | YES | YES | YES | YES | YES | YES | YES | YES |

Notes: This table reports the OLS estimation of our gravity model specification (i.e. without including non-linear adjustment and Heckman adjustment term). Dependent variable is year-over-year growth rate of direct cross-border / local affiliate stock of claims (expressed in log difference). Standard errors are clustered at country-pair level.