Online Appendix: Evolution of Bilateral Capital Flows to Developing Countries at Intensive and Extensive Margins

Juliana Araujo
International Monetary Fund

Povilas Lastauskas
University of Cambridge

Chris Papageorgiou
International Monetary Fund

Abstract

We report further robustness checks in this document. We analyze expanding network of capital flows and other data facts in Section 1. Section 2 covers copulae based methodology, which admits flexible specifications for unobservables in the extensive and intensive margin equations. Section 3 deals with negative values in flows whereas stability and crisis period are analyzed in Section 4. A persistence in extensive margin is explored in Section 5; a different variable for measuring lending rate, marginal product of capital, is used in Section 6; we also use size adjusted capital flows to explore if our results are consistent with the theory and robust to general expenditure generated heteroskedasticity, see Section 7. On the methodology side, we cover several aspects of economic model in Section 8 whereas econometrics is dealt with in Section 9. We also report on the details of computing intensive and extensive margins for the graphical exercise in the main text in Section 10. All data and codes are available upon request from the authors.

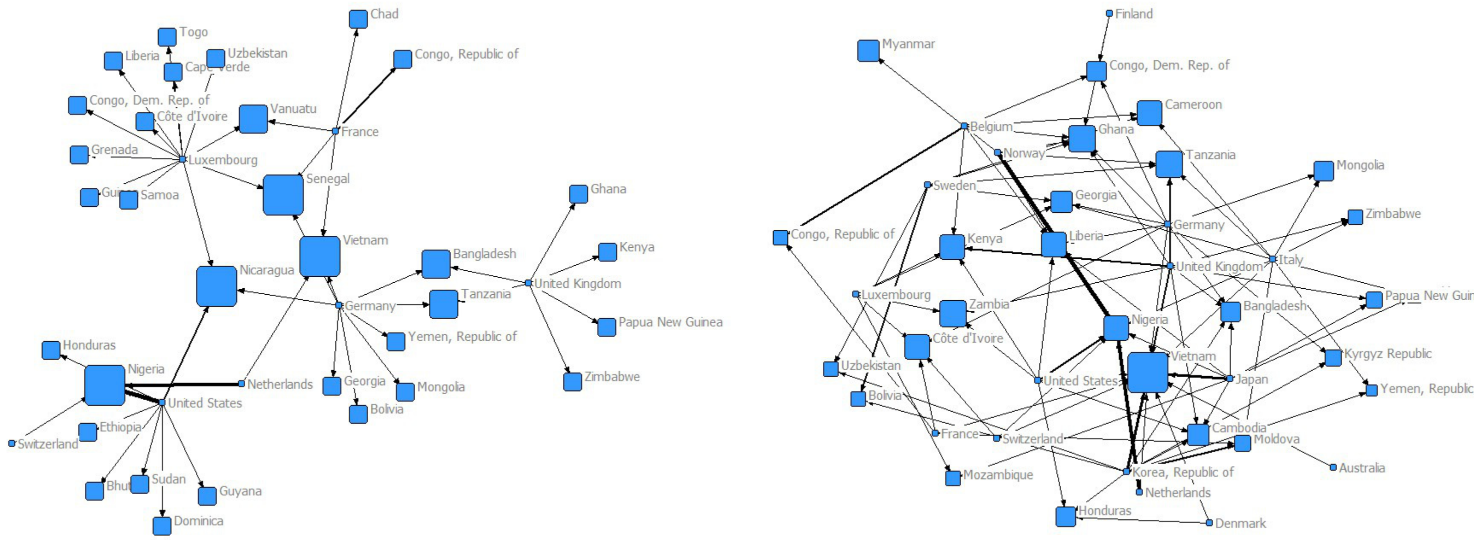

Fig. 1: Network Structure of Capital Flows to LICs in 2002 and 2010

Tab. 1: Aggregate Statistics of the Network of Capital Flows in 2002 and 2010 for LICs and MICs

	Average Degree	Weighted Average Degree	Core Membership
LIC 2002	7.335	21.736	Netherlands, Nigeria, United States
LIC 2010	9.819	103.046	Nigeria, Netherlands, Norway
MIC 2002	95.105	602.331	Mexico, United States
MIC 2010	30.597	2600.767	Brazil, China, India, Japan, Luxembourg, Switzerland, United States

Average Degree refers to the sum of edges of a vertex. Weighted Average Degree is the average of sum of weights of the edges of nodes.

1 Data Facts

1.1 Expanding Network of Capital Flows

We also analyze the changes in capital flows network. Figure 1 demonstrates how capital flows flew from advanced to developing countries in 2002 and 2010. In less than ten years, the network became way denser, with more investors, new receivers and more connections. This calls for further study of capital flow dynamics, in particular, to understand when the link is established and how much is being invested once it is in place. The cross-sectional dependence is also confirmed by a number of statistical tests.[1]

Table 1 collects information on aggregate statistics of capital network for both, LICs and MICs. Notice it is exactly for LICs that the average degree has increased - hence, the extensive margin played the most prominent role (the number of links attached to an economy increased). However, both extensive and intensive margins are taken into account once we compute the weighted average degree where the weights (volumes of capital flows) are considered. It is then clear that both LICs and MICs experienced substantial increases in capital flows, with a few economies receiving very large flows. Also, the core members have increased in number and changed identity. The emerging markets – such as Brazil, India and China (BIC) – emerged as core members in 2010 as compared to 2002.

[1] Available upon request.

1.2 Gravity-Type Relationship

Figure 2 demonstrates unconditional within variation of log FDI flows deviations from individual means against log of destination income deviations from individual means. There is a clear upward trend, thereby indicating of gravity forces - destination and source incomes being positively associated with FDI - at work.

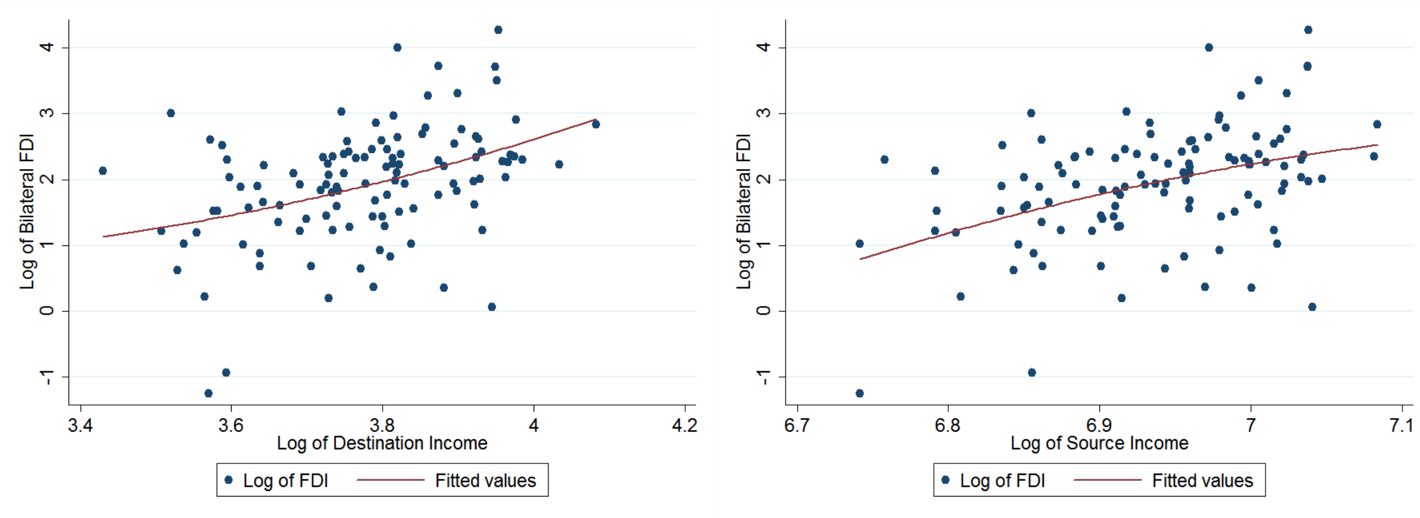

Fig. 2: Bilateral FDI vs Destination and Source Incomes in Overall Sample

It is clear that a failure to include directional fixed effects would have made positive correlation for both incomes. It seems that allowing for general equilibrium effects, subsumed within the fixed effects, make the net effect of destination income negative and in line with the findings in Baltagi, Egger, and Pfaffermayr (2008) and the theory in Bergstrand and Egger (2007).

2 Copulae-based Heckman Method

As a robustness check, we resort to the recent contribution by Hasebe and Vijverberg (2012). Parametric approach is criticized for its sensitivity to the distributional assumption. The alternative involves a copula, where a multivariate distribution is constructed from separately specified marginal distributions. The Generalized Tukey Lambda (GTL)-Copula approach inserts more flexible marginal distribution into the copula function. The GTL distribution is a versatile univariate distribution that permits a wide range of skewness of thick- or thin-tailed behavior in the data that it represents and, therefore, is a good candidate distribution for modeling the unobservables in the extensive and intensive margin equations. GTL-copula distribution effectively frees the extensive margin's model from any particular distributional assumption.

What is more, this flexibility is achieved with just a few additional parameters, which is both parsimonious and time-efficient relative to the semi- or non-parametric approaches. The collective set of copulas accommodates diverse dependence structures between two random variables. Unlike the traditional estimator (the normal-Gaussian estimator), the GTL-copula estimator is much less dependent on the presence of an instrument in the selection equation that fulfills the exclusion restriction, no longer making it problematic that the extensive margin equation contains the same explanatory variables as the capital volume equation.

Effectively, we deal with the flexible Heckman model, where we allow for the different distributions to be used, using a copula, in the intensive and extensive margins. In our application, the extensive margin is still a probit but an intensive margin contains a t-distribution, and all the errors are clustered at the

Tab. 2: Copula-based Capital Flows Models

	LIC	MIC	Overall Sample
Variables		Copulae	
Intensive Margin			
Log Destination Income	-1.489	1.443***	0.950**
	(1.326)	(0.482)	(0.459)
Log Source Income	0.347***	0.527***	0.525***
	(0.099)	(0.047)	(0.044)
Log Lending Rate	-0.689	-0.155	-0.262
	(0.655)	(0.199)	(0.189)
Log Labor Productivity	4.198	-0.130	0.578
	(2.619)	(0.960)	(0.902)
Extensive Margin			
Log Startup Costs	-0.102*	0.022	-0.024
	(0.061)	(0.050)	(0.038)
Log Lending Rate	-0.021	-0.146	-0.140*
	(0.185)	(0.092)	(0.084)
Log Labor productivity	1.184***	1.002***	0.992***
	(0.433)	(0.243)	(0.211)
Log σ	0.6549***	0.6554***	0.6266***
	(0.1178)	(0.0510)	(0.0440)
Ancillary θ	-0.6031	-0.8973***	-0.6599***
	(0.4194)	(0.1191)	(0.1149)

Note: time coverage is 2002-2010. θ refers to the dependence parameter in the copula framework. Robust standard errors in parentheses. *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term.

dyad level. Note that we control only for the destination fixed effects in the intensive margin for FDI and all directional fixed effects in the extensive margin.[2] Therefore, results should be treated with some care (and they should be compared to the Exponential Type II Tobit which is a less flexible counterpart of the below applied model).

Our main predictions, reported in Table 2, are in line with the previous results. It is worth mentioning that capital tends to flow from richer countries to countries with larger income (except for LICs), after conditioning on other theory implied variables. This is line with the theory and the largest difference from other regressions. Other variables follow previous findings: investor's income pins down the intensity of direct investment, whereas startup costs are negatively affecting only LICs' extensive margin. Labor productivity, as before, is crucial for attracting capital (in the copula case, however, result prevails in one, extensive, margin only).

3 Negative Values

In these sets of tables we introduce a method, developed in Eaton and Tamura (1994), where we estimate a new model, such that a dependent variable becomes $\ln(M_{in} + a)$, where $a = -\min\{M_{in}\}$+small number, thus taking negative values into account. The interpretation now changes since we lose all zeroes, and instead interpret values close to $\min\{M_{in}\}$ as reflecting no capital flow. Economic mechanisms generating negative values are potentially very different from the proposed theory but it is of interest to learn how

[2] This is because of convergence and multicollinearity issues. That is why we do not report portfolio flows results since, once fixed effects have been included, the Hessian was not negative semidefinite, which is needed for an optimum.

results change under this new setting.

3.1 FDI Regressions

Refer to Tables 3-5.

Tab. 3: Lognormal Hurdle for FDI

	LIC	MIC	Overall
	Lognormal Hurdle	Lognormal Hurdle	Lognormal Hurdle
Intensive Margin			
Log Destination GDP	-1.410	-1.540**	-1.533***
	(1.266)	(0.648)	(0.576)
Log Source GDP	2.369	5.297***	5.000***
	(1.468)	(0.883)	(0.766)
Log Lending Rate	-1.066**	-0.075	-0.155
	(0.445)	(0.155)	(0.150)
Log Labor Productivity	5.035***	2.438***	2.650***
	(1.772)	(0.770)	(0.727)
Extensive Margin			
Log Startup Costs	-0.200**	-0.262***	-0.228***
	(0.087)	(0.084)	(0.059)
Log Lending Rate	0.184	0.138	0.125
	(0.236)	(0.125)	(0.107)
Log Labor Productivity	1.139**	1.605***	1.426***
	(0.558)	(0.409)	(0.320)

Clustered standard errors for extensive margin in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tab. 4: Exponential Type 2 Tobit for FDI

	LIC ET2T	MIC ET2T	Overall ET2T
Intensive Margin			
Log Destination GDP	-2.096*	-1.082*	-1.191**
	(1.252)	(0.653)	(0.585)
Log Source GDP	1.964	4.674***	4.377***
	(1.483)	(0.884)	(0.775)
Log Lending Rate	-0.544	-0.184	-0.271
	(0.624)	(0.186)	(0.184)
Log Labor Productivity	6.542***	1.885**	2.421***
	(1.995)	(0.813)	(0.773)
Extensive Margin			
Log Startup Costs	-0.070	-0.159***	-0.122***
	(0.049)	(0.044)	(0.032)
Log Lending Rate	-0.307***	-0.140*	-0.252***
	(0.113)	(0.078)	(0.065)
Log Labor Productivity	-0.087	-0.326***	0.207***
	(0.093)	(0.100)	(0.052)

Clustered standard errors for extensive margin in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tab. 5: Mundlak-Chamberlain Panel for FDI

	LIC M-C Panel	MIC M-C Panel	Overall M-C Panel
Intensive Margin			
Log Destination GDP	-1.153	-1.284***	-1.182***
	(0.962)	(0.481)	(0.428)
Log Source GDP	3.608***	5.259***	4.989***
	(1.104)	(0.577)	(0.511)
Log Lending Rate	-1.122***	-0.230**	-0.297***
	(0.342)	(0.099)	(0.095)
Log Labor Productivity	3.519***	2.504***	2.509***
	(1.341)	(0.548)	(0.503)
Extensive Margin			
Log Startup Costs	-0.200**	-0.262***	-0.227***
	(0.086)	(0.084)	(0.058)
Log Lending Rate	0.150	0.156	0.140
	(0.244)	(0.117)	(0.105)
Log Labor Productivity	1.129**	1.619***	1.437***
	(0.502)	(0.358)	(0.283)

Clustered standard errors for extensive margin in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.2 Portfolio Flows

Due to limited data, we concentrate on the Mundlak-Chamberlain specification in the context of portfolio flows, see Table 6.

Tab. 6: Mundlak-Chamberlain Panel for Portfolio Investment

	LIC M-C Panel	MIC M-C Panel	Overall M-C Panel
Intensive Margin			
Log Contract Risk	0.425	0.463**	0.481**
	(1.410)	(0.235)	(0.232)
Log Destination GDP	0.290	0.070***	0.083***
	(0.235)	(0.026)	(0.026)
Log Source GDP	0.127	1.067***	1.024***
	(0.338)	(0.064)	(0.063)
Extensive Margin			
Log Contract Risk	-0.066	0.114	0.057
	(0.351)	(0.181)	(0.159)
Log Startup Costs	-0.305***	-0.212*	-0.238***
	(0.112)	(0.111)	(0.079)

Clustered standard errors for extensive margin in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4 Stability: Crisis Period

We explore parameter stability by including dummies for crisis period, 2008, not only as levels but also as interactions with all the already included variables. As reported in Table 7 and confirmed by a sequence of Chow tests (mentioned in the main text), there is no convincingly systematic evidence of visible change in the main parameter coefficients. We also reproduce tables for our preferred, correlated Mundlak-Chamberlain panel (under the heading of Corr. MC Panel), method. As is seen, crisis does not play an important role even when we account for all, intercept and slope, changes.

4.1 Simple Panel FDI and Portfolio Flows

Refer to Tables 7 and 8.

4.2 Mundlak-Chamberlain and Crisis: FDI

Though not reported, the cross-sectionally averaged source incomes turn out to be significant for intensive margin among all country groups. For the results on the margins, refer to Tables 9 and 10.

Tab. 7: Crisis and Panel Results of Bilateral FDI

	LIC			MIC			Overall		
	Pooled OLS	FE	RE (GLS)	Pooled OLS	FE	RE (GLS)	Pooled OLS	FE	RE (GLS)
Log Destination GDP	-0.927	-0.666	-0.920	-1.339**	-1.331**	-1.337***	-1.225**	-1.191**	-1.221***
	(0.863)	(0.370)	(0.997)	(0.569)	(0.523)	(0.487)	(0.481)	(0.446)	(0.434)
Log Source GDP	3.607***	4.185***	3.622***	5.310***	5.355***	5.313***	5.029***	5.120***	5.034***
	(1.249)	(1.118)	(1.175)	(0.763)	(0.907)	(0.593)	(0.656)	(0.721)	(0.528)
Log Lending Rate	-1.131***	-1.055***	-1.130***	-0.249**	-0.286**	-0.253**	-0.314***	-0.338***	-0.316***
	(0.376)	(0.183)	(0.351)	(0.109)	(0.114)	(0.102)	(0.107)	(0.089)	(0.098)
Log Labor Productivity	3.130**	2.134*	3.106**	2.553***	2.616***	2.558***	2.545***	2.520***	2.543***
	(1.321)	(0.993)	(1.386)	(0.645)	(0.542)	(0.554)	(0.576)	(0.445)	(0.507)
Log Dest. GDP x Crisis	-0.214	-0.254***	-0.215	0.034	0.031**	0.034	0.006	0.001	0.005
	(0.163)	(0.073)	(0.132)	(0.041)	(0.013)	(0.046)	(0.033)	(0.010)	(0.039)
Log Source GDP x Crisis	-0.103	-0.084	-0.102	0.026	0.026	0.026	0.004	0.004	0.004
	(0.092)	(0.054)	(0.098)	(0.042)	(0.031)	(0.043)	(0.038)	(0.030)	(0.039)
Log Lending x Crisis	0.011	-0.128	0.007	0.131	0.127***	0.131	0.122	0.100***	0.121
	(0.360)	(0.194)	(0.395)	(0.121)	(0.018)	(0.138)	(0.115)	(0.027)	(0.130)
Log Labor Productivity x Crisis	0.227	0.291**	0.229	0.083	0.098	0.083	0.082	0.103**	0.084
	(0.163)	(0.118)	(0.221)	(0.186)	(0.054)	(0.147)	(0.084)	(0.037)	(0.088)
Crisis	-0.403	-0.569	-0.409	-1.506	-1.606**	-1.506	-1.148	-1.262***	-1.156
	(2.191)	(1.118)	(2.059)	(1.999)	(0.503)	(1.586)	(0.876)	(0.328)	(0.931)
N	736	736	736	3209	3209	3209	3945	3945	3945

Robust standard errors for pooled OLS, Driscoll-Kraay standard errors for FE.Directional fixed effects (source and destination countries) as well as constants are included.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tab. 8: Crisis in Simple Panel Specifications

	Pooled OLS	FE	RE (GLS)
Log Destination GDP	0.068**	0.081***	0.078**
	(0.032)	(0.016)	(0.032)
Log Source GDP	0.998***	1.099***	1.048***
	(0.084)	(0.053)	(0.083)
Log Contract Risk	0.477*	0.419	0.468*
	(0.273)	(0.344)	(0.257)
Log Destination GDP x Crisis	-0.070	-0.087***	-0.079
	(0.056)	(0.023)	(0.053)
Log Source GDP x Crisis	0.099	0.043**	0.075
	(0.067)	(0.021)	(0.067)
Log Contract Risk x Crisis	-1.182	-0.457	-0.632
	(0.753)	(0.446)	(0.682)
Crisis	1.039	1.194**	0.879
	(1.127)	(0.531)	(1.048)
N	2738	2738	2738
R^2	0.671		

Robust standard errors for pooled OLS, Driscoll-Kraay standard errors for FE.

Directional fixed effects (source and destination countries) as well as constants are included.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tab. 9: Crisis and Mundlak-Chamberlain Panels for FDI: Intensive Margin

	LIC MC Panel	MIC MC Panel	Overall MC Panel
Intensive Margin			
Log Destination GDP	-0.992	-1.339***	-1.225***
	(0.976)	(0.485)	(0.432)
Log Source GDP	3.544***	5.313***	5.028***
	(1.152)	(0.592)	(0.526)
Log Lending Rate	-1.146***	-0.247**	-0.313***
	(0.343)	(0.102)	(0.097)
Log Labor Productivity	3.280**	2.551***	2.545***
	(1.358)	(0.552)	(0.505)
Log Destination GDP x Crisis	-0.202	0.034	0.006
	(0.129)	(0.046)	(0.039)
Log Source GDP x Crisis	-0.106	0.026	0.004
	(0.096)	(0.043)	(0.039)
Log Lending Rate x Crisis	0.035	0.131	0.122
	(0.387)	(0.138)	(0.130)
Log Labor Productivity x Crisis	0.222	0.081	0.082
	(0.217)	(0.147)	(0.088)
Crisis	-0.444	-1.491	-1.145
	(2.020)	(1.582)	(0.929)
N	736	3209	3945

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tab. 10: Crisis and Mundlak-Chamberlain Panels for FDI: Extensive Margin

	LIC MC Panel	MIC MC Panel	Overall MC Panel
Extensive Margin			
Log Startup Costs	-0.180**	0.059	-0.027
	(0.085)	(0.066)	(0.050)
Log Lending Rate	0.166	-0.101	-0.073
	(0.222)	(0.103)	(0.094)
Log Labor Productivity	0.910	0.858***	0.750***
	(0.555)	(0.275)	(0.241)
Log Startup Costs x Crisis	-0.061	0.042	0.026
	(0.092)	(0.059)	(0.048)
Log Lending Rate x Crisis	0.001	0.165	0.121
	(0.237)	(0.144)	(0.119)
Log Labor Productivity x Crisis	-0.182	0.153	0.090
	(0.126)	(0.123)	(0.066)
Crisis	1.781	-1.794	-1.065
	(1.643)	(1.310)	(0.824)
N	2375	5170	7800

Clustered standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.3 Mundlak-Chamberlain and Crisis: Portfolio

Refer to Tables 11 and 12.

5 Persistence in Extensive Margin

Provided fixed costs constituted a substantial part of total costs, we should observe a persistence in extensive margin. In other words, conditioning on the previous period's existence of an investment relationship, should affect decision for the current investment. Therefore, we use a dynamic probit to explore the temporal effects. If the margin at time zero is correlated with unobserved heterogeneity, inconsistent estimators are obtained (an example of the so-called initial conditions problem). In the Table 13, we report the first stages (probit) estimates after we allow for the investment probability to depend on the decision in the previous period. As opposed to Table 1 and Table 2 in the main text, here the extensive (not the intensive) margin is considered. The results indicate that persistence is found to be quite significant and robust across country groups.[3] However, having a short unbalanced panel, we stick to less computationally intensive methods whose results are testable in a more straightforward manner and which correspond to the static theory more closely.[4] We interpret the finding of persistence as suggesting that empirically fixed and startup costs need to be included in the static (steady state) specifications for the decision to invest (extensive margin).

[3] For the methodological issues on how to estimate a dynamic probit model, see Stewart (2006, 2007).

[4] Not only initial condition problem needs to be addressed but also multivariate normal probability functions of order T are required. One possibility is to resort to simulated likelihood method and estimate a simulated multivariate random effects probit. Relying on underlying assumptions and imposed structure, we feel a static specification is better suited for the data at hand.

Tab. 11: Crisis and Mundlak-Chamberlain Panels for Portfolio Investment: Intensive Margin

	LIC MC Panel	MIC MC Panel	Overall MC Panel
Intensive Margin			
Log Destination GDP	0.301	0.066**	0.077***
	(0.229)	(0.026)	(0.026)
Log Source GDP	0.361	1.076***	1.040***
	(0.345)	(0.064)	(0.063)
Log Contract Risk	1.417	0.441*	0.474**
	(1.442)	(0.234)	(0.231)
Log Destination GDP x Crisis	-0.150	-0.100	-0.078
	(0.189)	(0.063)	(0.053)
Log Source GDP x Crisis	0.306*	0.050	0.079
	(0.176)	(0.065)	(0.059)
Log Contract Risk x Crisis	-4.775	-0.017	-0.698
	(3.684)	(0.911)	(0.734)
Crisis	2.154	0.561	0.884
	(3.049)	(1.303)	(1.034)
N	147	2591	2738

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tab. 12: Crisis and Mundlak-Chamberlain Panels for Portfolio Investment: Extensive Margin

	LIC Corr. MC Panel	MIC Corr. MC Panel	Overall Corr. MC Panel
Extensive Margin			
Log Contract Risk	-0.157	-0.019	-0.014
	(0.254)	(0.113)	(0.102)
Log Startup Costs	-0.183**	0.012	-0.052
	(0.072)	(0.053)	(0.043)
Log Contract Risk x Crisis	-0.384	-0.047	-0.073
	(0.420)	(0.274)	(0.199)
Log Startup Costs x Crisis	-0.066	-0.002	0.208***
	(0.133)	(0.093)	(0.066)
Crisis	-0.114	-1.750***	-2.068***
	(0.796)	(0.421)	(0.326)
N	1117	4916	6035

Clustered standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tab. 13: Dynamic Probit Model for the Investment Decision

	LIC		MIC		Overall Sample	
Variables	No DFE	With DFE	No DFE	With DFE	No DFE	With DFE
Lagged Dummy for FDI	1.693***	0.958***	1.250***	0.632***	1.439***	0.752***
	(0.0677)	(0.0807)	(0.0405)	(0.0482)	(0.0341)	(0.0409)
Log Lending Rate	-0.144**	0.145	-0.0602	-0.0722	-0.105***	-0.0667
	(0.0697)	(0.284)	(0.0407)	(0.0996)	(0.0339)	(0.0940)
Log Labor productivity	-0.0416	0.468	-0.195***	0.783***	0.111***	0.528**
	(0.0532)	(0.534)	(0.0453)	(0.285)	(0.0265)	(0.244)
Log Startup Costs	-0.0362	-0.116	-0.0489**	0.128*	-0.0450***	0.0350
	(0.0315)	(0.0873)	(0.0209)	(0.0665)	(0.0169)	(0.0514)
Observations	2266		4609		6875	
Log likelihood	-971.63	-776.92	-2582.19	-2159.61	-3638.61	-2982.78
Pseudo R^2	0.27	0.36	0.17	0.31	0.2361	0.37

Note: time coverage is 2002-2010. DFE refers to the directional (source and destination) fixed effects. The estimation method is random effects dynamic probit model, estimated by the maximum likelihood. Integrals are evaluated using Gaussian–Hermite quadrature points. Standard errors in parentheses; *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term.

6 Panels with MPK

Tab. 14: Panel Results of Bilateral FDI Ignoring the Extensive Margin

	LIC			MIC			Overall Sample		
Variables	Pooled OLS	FE	RE (GLS)	Pooled OLS	FE	RE (GLS)	Pooled OLS	FE	RE (GLS)
Log Destination	0.84	1.25	0.86	-1.18*	-1.08***	-1.18**	-0.950*	-0.82***	-0.96**
Income	(1.516)	(1.272)	(1.506)	(0.609)	(0.322)	(0.520)	(0.555)	(0.279)	(0.483)
Log Source	2.49	2.61	2.53	5.25***	5.33***	5.25***	4.87***	4.92***	4.87***
Income	(1.636)	(2.036)	(1.617)	(0.836)	(0.432)	(0.654)	(0.748)	(0.395)	(0.603)
Log MPK	0.44	0.57	0.44	-0.41	-0.48	-0.41	-0.27	-0.33	-0.27
	(0.841)	(0.326)	(0.696)	(0.323)	(0.328)	(0.280)	(0.302)	(0.278)	(0.259)
Log Labor	0.63	-0.92	0.49	2.55***	2.37***	2.55***	2.46***	2.23***	2.47***
Productivity	(2.545)	(1.538)	(2.710)	(0.690)	(0.385)	(0.615)	(0.657)	(0.302)	(0.590)
Observations	499	499	499	2657	2657	2657	3156	3156	3156
Within R^2		0.11	0.11		0.20	0.20		0.19	0.19
Between R^2			0.61			0.69			0.70
Overall R^2			0.46			0.59			0.60

Note: time coverage is 2002-2010. Pooled OLS refer to pooled least squares where any time-specific effects are assumed to be fixed and the individual affects are centered around a common intercept; cluster robust standard errors are reported, also directional fixed effects (source and destination countries) are included. FE stands for fixed effects with the reported Driscoll-Kraay standard errors which help controlling for spatial (cross-sectional) dependence. RE (GLS) for random effects generalized least squares estimator (effectively a weighted average of within and between estimators which are not reported separately), also directional fixed effects (source and destination countries) are included. Standard errors in parentheses; *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term.

7 Size-adjusted Capital Flows

As pointed out by Santos-Silva and Tenreyro (2007), potential heteroskedasticity in the intensity equation may yield, inconsistency results due Jensen's inequality. In addition to Poisson results, as reported in the Section on Baseline Results of the main text, we further explore how results change if, instead of

Tab. 15: Bilateral FDI at Extensive and Intensive Margins with MPK

	LIC			MIC			Overall Sample		
Variables	Pooled Lognormal Hurdle	Pooled Exp. Type II Tobit	Mundlak Chamberlain Panel	Pooled Lognormal Hurdle	Pooled Exp. Type II Tobit	Mundlak Chamberlain Panel	Pooled Lognormal Hurdle	Pooled Exp. Type II Tobit	Mundlak Chamberlain Panel
Intensive Margin of FDI									
Log Destination Income	1.20	0.654	1.04	-1.90**	-1.67**	-1.19**	-1.59**	-1.32*	-0.795
	(1.81)	(1.951)	(1.576)	(0.765)	(0.795)	(0.551)	(0.694)	(0.727)	(0.513)
Log Source Income	1.23	1.045	1.46	5.25***	5.01***	5.34***	4.73***	4.38***	4.680***
	(1.89)	(2.095)	(1.66)	(0.998)	(1.046)	(0.696)	(0.883)	(0.933)	(0.639)
Log MPK	0.39	-.684	-.246	-0.33	-0.472	0.39	-0.18	-0.34	0.348
	(0.91)	(1.029)	(0.802)	(0.37)	(0.403)	(0.441)	(0.347)	(0.373)	(0.376)
Log Labor productivity	1.80	1.997	0.390	3.66***	3.15***	3.06***	3.57***	3.14***	2.83***
	(3.31)	(3.850)	(3.344)	(0.89)	(0.971)	(0.717)	(0.849)	(0.940)	(0.683)
Extensive Margin of FDI									
Log Startup Costs	-0.0003	-.0003	-0.0003	-0.0006	-.0006	-0.0006	-0.0008	-.0008	-0.0008
	(0.0009)	(0.0010)	(0.0009)	(0.0022)	(0.0022)	(0.0022)	(0.0008)	(0.0008)	(0.0008)
Log MPK	7.090	6.966	7.090	10.36***	10.357***	10.36***	10.41***	10.36***	10.41***
	(6.4386)	(6.342)	(6.4386)	(3.0960)	(3.091)	(3.0960)	(2.76)	(2.757)	(2.76)
Log Labor productivity	0.0007***	0.0007***	0.0007***	0.00006***	0.00006***	0.00006***	0.00007***	0.00007***	0.00007***
	(0.0002)	(0.0002)	(0.0002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(.00002)	(0.00002)
$\hat{\sigma}$ ($\hat{\sigma_u}$/ $\hat{\sigma_\varepsilon}$)	1.53	1.83	.989/1.17	1.65	1.64	1.18/1.14	1.65	1.64	1.17/1.14
	(0.052)	(0.177)	(0.091/0.046)	(0.027)	(0.028)	(0.047/0.018)	(0.025)	(.026)	(0.042/0.017)
$\hat{\rho}$	-	-0.74***	0.418	-	0.075	0.52***	-	0.11*	0.51***
		(0.131)	(0.052)		(0.055)	(0.022)		(0.06)	(0.02)
Observations (positive/all)	499/1211	460/1483	458/1211	2657/4050	2447/4050	2447/4050	3156/5533	2907/5533	2907/5533
Log likelihood	-3939.58	-3860.76	-3817.86	-20994.15	-20570.64	-20126.67	-25014.95	-24514.76	-24011.79

Note: time coverage is 2002-2010 The difference in observations in ET2T arises because all negative flows are excluded whereas in other estimations they are truncated as zeroes, potentially carrying information on the existence of a link but absence of a positive capital flow. Mundlak-Chamberlain method includes time averages of relevant explanatory variables and inverse Mills ratios from the first stage Probit regressions. Time averages are omitted in the first stage due to collinearity. Robust standard errors in parentheses. *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term and directional fixed effects (source and destination countries).

Tab. 16: Lognormal Hurdle for Size-adjusted FDI

	LIC Lognormal Hurdle	MIC Lognormal Hurdle	Overall Lognormal Hurdle
Intensive Margin			
Log Lending Rate	-0.323	-0.153	-0.156
	(0.431)	(0.159)	(0.154)
Log Labor Productivity	2.636***	1.589***	1.625***
	(0.820)	(0.495)	(0.453)
Extensive Margin			
Log Startup Costs	-0.189**	0.061	-0.029
	(0.085)	(0.066)	(0.049)
Log Lending Rate	0.153	-0.073	-0.060
	(0.220)	(0.101)	(0.093)
Log Labor Productivity	0.894	1.036***	0.882***
	(0.551)	(0.272)	(0.238)

Clustered standard errors for extensive margin in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

regressing capital flows on all the theory implied variables, we first divide them by source and destination income. These variables are purported to be the main channels of heteroskedasticity. This procedure helps to alleviate heteroskedasticity, and, by that argument, should be more reliable. Results are reported in Tables 16-19.

Main results

To ease comparison with the results in the main paper, we report main tables from the paper. Refer to Tables 20 and 21.

Tab. 17: Exponential Type 2 Tobit for Size-adjusted FDI

	LIC ET2T	MIC ET2T	Overall ET2T
Intensive Margin			
Log Lending Rate	-0.669	-0.162	-0.234
	(0.556)	(0.186)	(0.181)
Log Labor Productivity	2.155**	0.945***	1.092***
	(0.853)	(0.366)	(0.350)
Extensive Margin			
Log Startup Costs	-0.073	-0.159***	-0.122***
	(0.051)	(0.044)	(0.032)
Log Lending Rate	-0.312***	-0.141*	-0.252***
	(0.113)	(0.079)	(0.065)
Log Labor Productivity	-0.104	-0.325***	0.207***
	(0.092)	(0.100)	(0.052)

Clustered standard errors for extensive margin in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tab. 18: Mundlak-Chamberlain Panel for Size-adjusted FDI

	LIC M-C Panel	MIC M-C Panel	Overall M-C Panel
Intensive Margin			
Log Lending Rate	-1.063***	-0.274***	-0.337***
	(0.331)	(0.097)	(0.093)
Log Labor Productivity	1.743***	1.727***	1.714***
	(0.588)	(0.209)	(0.197)
Extensive Margin			
Log Startup Costs	-0.188**	0.061	-0.028
	(0.085)	(0.066)	(0.049)
Log Lending Rate	0.151	-0.061	-0.047
	(0.220)	(0.101)	(0.093)
Log Labor Productivity	0.893	1.040***	0.888***
	(0.550)	(0.273)	(0.238)

Clustered standard errors for extensive margin in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tab. 19: Mundlak-Chamberlain Panel for Size-adjusted Portfolio Investment

	LIC M-C Panel	MIC M-C Panel	Overall M-C Panel
Intensive Margin			
Log Contract Risk	0.214	0.429***	0.393***
	(0.334)	(0.138)	(0.128)
Extensive Margin			
Log Contract Risk	-0.176	0.105	0.081
	(0.257)	(0.106)	(0.099)
Log Startup Costs	-0.117*	0.329***	0.208***
	(0.070)	(0.048)	(0.042)

Clustered standard errors for extensive margin in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

8 Theoretical Derivations

8.1 Firms' Decisions

Firms take the residual demand in (2) as given and maximize profits by choosing a price. Omitting a time subscript for simplicity, note that in any country i, the profit for a $\varphi-$firm that might have entered market n is given by

$$\max_{\{p_i(\omega)\}\geq 0} \pi_i\left(\varphi\right) = \left(p_i\left(\omega\right) - v_i \tau_{in}\varphi\right)\left(\left(Q_i + I_i\right)\left(\frac{p_i\left(\omega\right)}{P_i}\right)^{-\sigma}\right) - v_n f_{in}, \tag{1}$$

where the first order condition leads to

$$\left(1 - \sigma + \sigma v_i \tau_{in}\varphi p_i\left(\omega\right)^{-1}\right)\left(\left(Q_i + I_i\right)\left(\frac{p_i(\omega)}{P_i}\right)^{-\sigma}\right) = 0,$$

and produces the price $p_{in}\left(\omega\right) = \frac{\sigma}{\sigma-1} v_i \tau_{in}\varphi$. Notice that the revenue is given by

$$\begin{aligned} R_{in}\left(\varphi\right) &\equiv p_{in}\left(\omega\right)\left(q_{in}\left(\omega\right) + i_{in}\left(\omega\right)\right) = p_{in}\left(\omega\right)^{1-\sigma}\left(Q_n + I_n\right) P_n^{\sigma} \\ &= \left(\tfrac{\sigma}{\sigma-1} c_{in}\left(\varphi\right)\right)^{1-\sigma}\left(Q_n + I_n\right) P_n^{\sigma} = \left(\tfrac{\sigma}{\sigma-1} c_{in}\left(\varphi\right)\right)^{1-\sigma} \tfrac{P_n\left(Q_n + I_n\right)}{P_n^{1-\sigma}}. \end{aligned}$$

Plugging price into profit yields $\pi_{in}\left(\varphi\right) = \frac{p_{in}(\varphi)(q_{in}(\varphi)+i_{in}(\varphi))}{\sigma} - v_n f_{in} = \frac{R_{in}(\varphi)}{\sigma} - v_n f_{in}$. Finally, the payment to the factors of production are uncovered from the cost minimization problem,

$$\begin{aligned} &\min\nolimits_{\{l_i(\varphi)\}\geq 0, \{k_i(\varphi)\}\geq 0} r_i k\left(\varphi\right) + w_i l\left(\varphi\right) = C_{in}\left(q\right) = \tfrac{r_i^{\alpha_i} w_i^{1-\alpha_i}}{\alpha_i^{\alpha_i}\left(1-\alpha_i\right)^{1-\alpha_i}}\left(q_i + i_i\right)\tau_{in}\varphi \\ &\text{s.t. } y = \tfrac{k^{\alpha_i}(Al)^{1-\alpha_i}}{\varphi} = q_i\left(\omega\right) + i_i\left(\omega\right) = \left(Q_i + I_i\right)\left(\tfrac{p_i(\omega')}{P_i}\right)^{-\sigma}. \end{aligned}$$

The associated Lagrangian yields the results as in the main text. However, a simpler approach is to observe that with the competitive factor markets, the payments are just marginal costs

$$\begin{aligned} w_i l_i\left(\varphi\right) &= \left(1 - \alpha_i\right) p_{in}\left(\varphi\right) y_{in}\left(\varphi\right) = \left(1 - \alpha_i\right) R_{in}\left(\varphi\right), \\ r_i k_i\left(\varphi\right) &= \alpha_i p_{in}\left(\varphi\right) y_{in}\left(\varphi\right) = \alpha_i R_{in}\left(\varphi\right). \end{aligned}$$

As stated in the main text, $w_i l_i + r_i k_i = \frac{\sigma-1}{\sigma} E_i$ where a wedge between total revenues and factor remuneration is caused by monopolistic competition in the goods market.

Tab. 20: Panel Bilateral FDI at Extensive and Intensive Margins

	LIC			MIC			Overall Sample		
Variables	Lognormal Hurdle	Exponential Type II Tobit	Mundlak Chamberlain Panel	Lognormal Hurdle	Exponential Type II Tobit	Mundlak Chamberlain Panel	Lognormal Hurdle	Exponential Type II Tobit	Mundlak Chamberlain Panel
Intensive Margin of Log FDI									
Log Destination GDP	-1.41	-1.872	-1.153	-1.54**	-1.143*	-1.284***	-1.53***	-1.230**	-1.182***
	(1.266)	(1.262)	(-0.962)	(0.648)	(0.65)	(0.481)	(0.576)	(0.583)	(0.428)
Log Source GDP	2.37	1.965	3.608***	5.30***	4.673***	5.259***	5.00***	4.373***	4.989***
	(1.468)	(1.484)	(1.104)	(0.883)	(0.885)	(0.577)	(0.766)	(0.776)	(0.511)
Log Lending Rate	-1.07**	-0.992*	-1.122***	-0.07	-0.17	-0.230**	-0.156	-0.243	-0.297***
	(0.445)	(0.542)	(0.342)	(0.156)	(0.187)	(0.0993)	(0.150)	(0.181)	(0.0954)
Log Labor productivity	5.03***	6.582***	3.519***	2.44***	2.047**	2.504***	2.650***	2.522***	2.509***
	(1.772)	(2.038)	(1.341)	(0.770)	(0.805)	(0.548)	(0.727)	(0.773)	(0.503)
Extensive Margin of FDI									
Log Startup Costs	-0.189**	-0.185**	-0.188**	0.0614	0.0586	0.0613	-0.0287	-0.0305	-0.0281
	(0.0854)	(0.0886)	(0.0854)	(0.0658)	(0.0657)	(0.0659)	(0.0493)	(0.0493)	(0.0493)
Log Lending Rate	0.153	0.154	0.1506	-0.0726	-0.0722	-0.0611	-0.0596	-0.0574	-0.0465
	(0.22)	(0.22)	(0.2205)	(0.101)	(0.101)	(0.101)	(0.0932)	(0.0931)	(0.093)
Log Labor productivity	0.894	0.92	0.893	1.036***	1.041***	1.040***	0.882***	0.891***	0.888***
	(0.551)	(0.566)	(0.55)	(0.272)	(0.273)	(0.273)	(0.238)	(0.238)	(0.238)
Observations			687/2375			2962/5170			3649/7800
Log likelihood		-4748.95	-4650.88		-22304.34	-21720.51		-27168.72	-26459.18

Note: time coverage is 2002-2010 The difference in observations in ET2T arises because all negative flows are excluded whereas in other estimations they are truncated as zeroes, potentially carrying information on the existence of a link but absence of a positive capital flow. Mundlak Chamberlain method includes time averages of relevant explanatory variable. Time averages of startup costs and labor productivity are omitted in the first stage due to collinearity. Clustered standard errors in parentheses. *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term and directional fixed effects (source and destination countries).

Tab. 21: Panel Bilateral Portfolio Investment at Extensive and Intensive Margins

	LIC	MIC	Overall Sample	
Variables	Lognormal Hurdle	Lognormal Hurdle	Lognormal Hurdle	Mundlak Chamberlain Panel
Intensive Margin of Log Portfolio				
Log Destination GDP	0.290	0.059*	0.073**	0.0825***
	(0.235)	(0.032)	(0.032)	(0.026)
Log Source GDP	0.127	1.034***	0.983***	1.024***
	(0.338)	(0.085)	(0.082)	(0.063)
Log Contract Risk	0.425	0.469**	0.480*	0.481**
	(1.410)	(0.275)	(0.269)	(0.232)
Extensive Margin of Portfolio				
Log Startup Costs	-0.117*	0.329***	0.208***	0.208***
	(0.070)	(0.048)	(0.0417)	(0.0417)
Log Contract Risk	-0.176	0.105	0.0807	0.0807
	(0.257)	(0.106)	(0.0987)	(0.0987)
Observations	147	2591	2738	4932

Note: time coverage is 2002-2010. Clustered standard errors in parentheses. *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term and directional fixed effects (source and destination countries). ET2T extensive margin is the same as that of Normal hurdle due to collinearity of time averages of startup costs and contract risk.

8.2 Agents' Decisions

The utility maximization with respect to the consumption and investment good $q_{it}+i_{it}$ (think of it as an amalgam of two goods, both valued by the agent) ,

$$\ln\left(Q_{it}+I_{it}\right)=\ln\left(\int_{\omega\in\Omega_i}\left(q_{in,t}\left(\omega\right)+i_{in,t}\left(\omega\right)\right)^{\frac{\sigma-1}{\sigma}}d\omega\right)^{\frac{\sigma}{\sigma-1}}$$

every period subject to $R_{it}=\sum_n\int_{\omega\in\Omega_i}p_{in,t}\left(\omega\right)\left(q_{in,t}\left(\omega\right)+i_{in,t}\left(\omega\right)\right)d\omega=\sum_n M_{in,t}$, which is taken as given, yields the value of demand

$$M_{in,t}=\left(\frac{p_{it}\tau_{in}}{P_{nt}}\right)^{1-\sigma}R_{nt}, \tag{2}$$

since income is equal to expenditure in equilibrium. Note that this functional form will remain intact despite the period – demand will be a function of expenditure. Firms take the residual demand in (2) as given and maximize profits by choosing a price. In any country n, the price of a variety ω from country i is given by

$$\begin{aligned}p_{in,t}\left(\omega\right)&=\tfrac{\sigma}{\sigma-1}\tau_{in}v_{it}\varphi\\&=\tau_{in}p_{it}\left(\omega\right)\end{aligned}\quad\text{for all }\omega\in\Omega_{in,t}. \tag{3}$$

Note that $p_{in,t}$ depends on trade costs τ_{in}, which could be interpreted as 'iceberg costs'. Hence, for unit to arrive, τ_{in} more product shall be sent as this share 'melts' on the way.

8.3 Extensive Margin

Refer to the definition of the latent variable Z_{in}^{FDI} – in our case,

$$\begin{aligned}Z_{in}^{FDI}&=\frac{\frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma}R_n(v_n\varphi_L)^{1-\sigma}P_n^{\sigma-1}}{v_nf_{in}^{FDI}}\\&=\frac{R_nR_i\varphi_L^{1-\sigma}}{\sigma K_iv_n^\sigma f_{in}^{FDI}}\left(P_n\Pi_i\right)^{\sigma-1}=\frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma}\frac{P_n^{\sigma-1}R_n\varphi_L^{1-\sigma}}{v_n^\sigma f_{in}^{FDI}}.\end{aligned}$$

As stated in the main text, the separation of input cost does not work in our model, as opposed to Behar and Nelson (2009), because our model relies on destination country's labor hiring and capital rental rates. Further, observe that scaling $\mathsf{z}_{in,t}\equiv\ln Z_{in,t}^{FDI}$ by the standard deviation σ_η, we obtain the probability ρ_{in} that i decides to invest to n, conditional on the observed variables, $\rho_{in,t}=\Pr\left(\mathsf{z}_{in,t}^{\star}>0\mid\text{observed variables}\right)$. This helps identifying a probability of establishing a link between two economies. Note that our setup can be easily adapted for trade flows too.

The least productive firm from country i sells in country $\left(\varphi_{in,t}^{EXP}\tau_{in}\right)^{1-\sigma}\frac{R_{nt}}{\sigma^\sigma(\sigma-1)^{1-\sigma}P_{nt}^{1-\sigma}}=v_{it}^\sigma f_{in,t}^{EXP}$. Note that the condition for ordering of the cutoffs between FDI and exporting is as follows (omitting time variable as the conditions hold in all periods): $\varphi_{in}^{FDI}<\varphi_{in}^{EXP}$, leading to $\varphi_{in}^{EXP}=\frac{1}{\tau_{in}}\left[\frac{v_i^\sigma f_{in}^{EXP}\sigma^\sigma(\sigma-1)^{1-\sigma}P_n^{1-\sigma}}{R_n}\right]^{1/(1-\sigma)}$ and $\varphi_{in}^{FDI}=\left[\frac{v_n^\sigma f_{in}^{FDI}\sigma^\sigma(\sigma-1)^{1-\sigma}P_n^{1-\sigma}}{R_n}\right]^{1/(1-\sigma)}$. Therefore, the requirement is equivalent to $\tau_{in}^{\sigma-1}f_{in}^{EXP}<\left(\frac{v_n}{v_i}\right)^\sigma f_{in}^{FDI}$, where the outcome is determined by relative sizes of fixed costs, transportation costs, and relative costs of capital and labor as embodied in v. We do not model trade flows in the current paper, so leave this aspect not further developed.

8.4 Optimal Portfolio Choice

As mentioned in the main text, we follow Tille and van Wincoop (2010) and Devereux and Sutherland (2011) to derive the optimal weights of portfolio investment. Before delving into the optimal weights, note that the optimization exercise, faced by agents, can be described by the following:

$$\begin{aligned}
\mathcal{L}_i = & \quad \ln\left(Q_{i1}+I_{i1}\right)+\frac{1}{1+\rho}\mathbb{E}_1\ln\left(Q_{i2}+I_{i2}\right)-\lambda\left[P_{i2}\left(Q_{i2}+I_{i2}\right)-R_i^p\left(E_{i1}-P_{i1}\left(Q_{i1}+I_{i1}\right)\right)\right], \\
\frac{\partial\mathcal{L}}{\partial(Q_{i1}+I_{i1})} = & \quad \frac{1}{Q_{i1}+I_{i1}}-\lambda R_i^p P_{i1}=0, \\
\frac{\partial\mathcal{L}}{\partial(Q_{i2}+I_{i2})} = & \quad \frac{1}{1+\rho}\mathbb{E}_1\frac{1}{Q_{i2}+I_{i2}}-\lambda=0,
\end{aligned}$$

where $P_{i2}=1$. The constraint is binding, so the Lagrange multiplier is not trivial: substitution one equation into another yields the Euler equation, reported in the main text. The remaining part is to pin down investment weights in R_i^p, provided that $\sum_{n=1}^{\mathbb{N}} w_{in}=1$, and taking world expenditures as given. We will use a Taylor expansion for the exponential function. Recall that a stochastic return can be decomposed into its components of different orders, $R_i=R_i(0)+R_i(1)+R_i(2)+\ldots$. The first order accounts for the standard deviation, the second order includes variance, covariance, square of error or a product of two errors.

Before proceeding to the solution method, we define (first order components of) stochastic returns as $\ln R_n(1)=\ln R(0)\left(\varepsilon_n+\theta_n\varepsilon_g\right)$ and $\ln R_g(1)=\ln R(0)\,\theta_g\varepsilon_g$, where ε_n is a an economy n payoff innovation, ε_g is the global payoff shock, θ_n and θ_g measure the strength of transmission of global shock on economy-specific and global returns, respectively.

It is then easy to see that a linear Taylor expansion of e^R is

$$e^{R(0)}+e^{R(0)}\left(R_i-R(0)\right)=e^{R(0)}\left(1+R_i-R(0)\right).$$

Using definition of R_i as the sum of its different ordered terms, we obtain $e^{R(0)}R(1)$. As for the second order term, start with the quadratic Taylor expansion of e^R:

$$\begin{aligned}
&e^{R(0)}+e^{R(0)}\left(R_i-R(0)\right)+0.5e^{R(0)}\left(R_i-R(0)\right)^2 \\
&=e^{R(0)}\left(1+\left(R_i-R(0)\right)+0.5\left(R_i-R(0)\right)^2\right).
\end{aligned}$$

Plugging in the definition of R_i, it follows that the second order component is given by

$$e^{R(0)}\left(R(2)+0.5\left(R(1)\right)^2\right).$$

Having these tools, let us return to the first order conditions. As stated, the following portfolio Euler equations must hold. Let's start with the zero-order terms first. Exponentiation yields

$$\begin{aligned}
\mathbb{E}_1 e^{-\ln Q_i+\ln R_n} &= \mathbb{E}_1 e^{-\ln Q_i+\ln R_f}, \\
\mathbb{E}_1 e^{-\ln Q_i+\ln R_g} &= \mathbb{E}_1 e^{-\ln Q_i+\ln R_f}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbb{E}_1 e^{-\ln Q_i(0)+\ln R_n(0)} &= \mathbb{E}_1 e^{-\ln Q_i(0)+\ln R_f(0)} \\
\mathbb{E}_1 e^{-\ln Q_i(0)+\ln R_g(0)} &= \mathbb{E}_1 e^{-\ln Q_i(0)+\ln R_f(0)},
\end{aligned}$$

giving $R_n(0)=R_f(0)$ and $R_g(0)=R_f(0)$. The first order component is $\left(e^{x(0)}x(1)\right)$

$$\begin{aligned}
e^{-\ln Q_i(0)+\ln R_n(0)}\mathbb{E}_1\left(-\ln Q_{i2}(1)+\ln R_{n2}(1)\right) &= e^{-\ln Q_i(0)+\ln R_f(0)}\mathbb{E}_1\left(-\ln Q_{i2}(1)+\ln R_{f2}(1)\right) \\
e^{-\ln Q_i(0)+\ln R_g(0)}\mathbb{E}_1\left(-\ln Q_{i2}(1)+\ln R_{g2}(1)\right) &= e^{-\ln Q_i(0)+\ln R_f(0)}\mathbb{E}_1\left(-\ln Q_{i2}(1)+\ln R_{f2}(1)\right).
\end{aligned}$$

Since $R_n(0)=R_f(0)$ and $R_g(0)=R_f(0)$,

$$\begin{aligned}
\mathbb{E}_1\ln R_{n2}(1) &= \mathbb{E}_1\ln R_{f2}(1) \\
\mathbb{E}_1\ln R_{g2}(1) &= \mathbb{E}_1\ln R_{f2}(1).
\end{aligned}$$

Finally, the second order components for the two conditions are

$$e^{-\ln Q_i(0)+\ln R_n(0)}\mathbb{E}_1\left(-\ln Q_{i2}(2)+\ln R_{n2}(2)+0.5\mathbb{E}_1\left(-\ln Q_{i2}(1)+\ln R_{n2}(1)\right)^2\right)$$
$$=e^{-\ln Q_i(0)+\ln R_f(0)}\mathbb{E}_1\left(-\ln Q_{i2}(2)+\ln R_{f2}(2)+0.5\mathbb{E}_1\left(-\ln Q_{i2}(1)+\ln R_{f2}(1)\right)^2\right),$$
$$e^{-\ln Q_i(0)+\ln R_g(0)}\mathbb{E}_1\left(-\ln Q_{i2}(2)+\ln R_{g2}(2)+0.5\mathbb{E}_1\left(-\ln Q_{i2}(1)+\ln R_{g2}(1)\right)^2\right)$$
$$=e^{-\ln Q_i(0)+\ln R_f(0)}\mathbb{E}_1\left(-\ln Q_{i2}(2)+\ln R_{f2}(2)+0.5\mathbb{E}_1\left(-\ln Q_{i2}(1)+\ln R_{f2}(1)\right)^2\right).$$

The difficulty arises due to the quadratic term expectation of which is not equal to the expectation squared. However, we can rewrite

$$\mathbb{E}_1\ln R_{n2}(2)+0.5\mathbb{E}_1\left(-\ln Q_{i2}(1)+\ln R_{n2}(1)\right)^2=\mathbb{E}_1\ln R_{f2}(2)+0.5\mathbb{E}_1\left(-\ln Q_{i2}(1)+\ln R_{f2}(1)\right)^2,$$
$$\mathbb{E}_1\ln R_{g2}(2)+0.5\mathbb{E}_1\left(-\ln Q_{i2}(1)+\ln R_{g2}(1)\right)^2=\mathbb{E}_1\ln R_{f2}(2)+0.5\mathbb{E}_1\left(-\ln Q_{i2}(1)+\ln R_{f2}(1)\right)^2.$$

Using the fact that $\ln R_{f2}(1)=0$, and $\mathbb{E}_1\ln R_{n2}(1)=\mathbb{E}_1\ln R_{f2}(1)$, $\mathbb{E}_1\ln R_{g2}(1)=\mathbb{E}_1\ln R_{f2}(1)$, we obtain

$$\mathbb{E}_1\left(\ln R_{n2}(2)-\ln R_{f2}(2)\right)=0.5\mathbb{E}_1\left(-\ln Q_{i2}(1)\right)^2-0.5\mathbb{E}_1\left(\ln R_{n2}(1)-\ln Q_{i2}(1)\right)^2,$$
$$\mathbb{E}_1\left(\ln R_{g2}(2)-\ln R_{f2}(2)\right)=0.5\mathbb{E}_1\left(-\ln Q_{i2}(1)\right)^2-0.5\mathbb{E}_1\left(\ln R_{g2}(1)-\ln Q_{i2}(1)\right)^2.$$

These can be simplified to obtain

$$\mathbb{E}_1\left(\ln R_{n2}(2)-\ln R_{f2}(2)\right)=\mathbb{E}_1\ln R_{n2}(1)\ln Q_{i2}(1)-0.5\mathbb{E}_1\ln R_{n2}(1)\ln R_{n2}(1),$$
$$\mathbb{E}_1\left(\ln R_{g2}(2)-\ln R_{f2}(2)\right)=\mathbb{E}_1\ln R_{g2}(1)\ln Q_{i2}(1)-0.5\mathbb{E}_1\ln R_{g2}(1)\ln R_{g2}(1).$$

Recalling that returns are driven by stochastic shocks, write $\ln R_n(1)=\ln R(0)(\varepsilon_n+\theta_n\varepsilon_g)$ and $\ln R_g(1)=\ln R(0)\theta_g\varepsilon_g$. Therefore, $\text{Var}R_{n2}(1)=\ln R(0)^2\left(\sigma_n^2+\theta_n^2\sigma_g^2\right)$, $\text{Var}R_g(1)=\ln R(0)^2\theta_g^2\sigma_g^2$, and $\mathbb{E}_1\ln R_n(1)\ln R_n(1)=\ln R(0)^2\left(\sigma_n^2+\theta_n^2\sigma_g^2\right)$, $\mathbb{E}_1\ln R_n(1)\ln R_g(1)=\ln R(0)^2\theta_n\theta_g\sigma_g^2$, $\mathbb{E}_1\ln R_g(1)\ln R_g(1)=\ln R(0)^2\theta_g^2\sigma_g^2$. This gives

$$\mathbb{E}_1\left(\ln R_{n2}(2)-\ln R_{f2}(2)\right)=\mathbb{E}_1\ln R_{n2}(1)\ln Q_{i2}(1)-0.5\ln R(0)^2\left(\sigma_n^2+\theta_n^2\sigma_g^2\right),$$
$$\mathbb{E}_1\left(\ln R_{g2}(2)-\ln R_{f2}(2)\right)=\mathbb{E}_1\ln R_{g2}(1)\ln Q_{i2}(1)-0.5\ln R(0)^2\theta_g^2\sigma_g^2.$$

We need to get the first order component of consumption from $Q_{i2}=W_{i1}R_{i2}^p$,

$$e^{\ln Q_{i2}(0)}=e^{\ln W_{i1}(0)+\ln R_{i2}^p(0)},$$

and

$$e^{\ln Q_{i2}(0)}\left(\ln Q_{i2}(1)\right)=e^{\ln W_{i1}(0)+\ln R_{i2}^p(0)}\left(\ln W_{i1}(1)+\ln R_{i2}^p(1)\right),$$

implying $\ln Q_{i2}(1)=\ln W_{i1}(1)R_{i2}^p(1)$, where

$$e^{\ln R_{i2}^p(0)}=\sum_{n=1}^{\mathbb{N}}w_{in}e^{\ln R_n(0)}+w_{ig}e^{\ln R_g(0)}+w_{if}e^{\ln R_f(0)},$$
$$e^{\ln R_{i2}^p(0)}\ln R_{i2}^p(1)$$
$$=\left(\sum_{n=1}^{\mathbb{N}}w_{in}e^{\ln R_n(0)}+w_{ig}e^{\ln R_g(0)}+w_{if}R_f(0)\right)\left(\sum_{n=1}^{\mathbb{N}}w_{in}\ln R_n(1)+w_{ig}\ln R_g(1)+w_{if}\ln R_f(1)\right),$$

or $\ln R_{i2}^p(1)=\sum_{n=1}^{\mathbb{N}}w_{in}\ln R_n(1)+w_{ig}\ln R_g(1)+w_{if}\ln R_f(1)$. Hence,

$$\mathbb{E}_1\left(\ln R_n(2)-\ln R_f(2)\right)=\mathbb{E}_1\ln R_{n2}(1)\ln W_{i1}(1)+\sum_{n=1}^{\mathbb{N}}w_{in}\mathbb{E}_1\ln R_n(1)\ln R_n(1)$$
$$+w_{ig}\mathbb{E}_1\ln R_n(1)\ln R_g(1)+w_{if}\mathbb{E}_1\ln R_n(1)\ln R_f(1)-0.5\ln R(0)^2\left(\sigma_n^2+\theta_n^2\sigma_g^2\right)$$
$$=\mathbb{E}_1\ln R_n(1)\ln W_{i1}(1)+\ln R(0)^2\tau_{in}w_{in}\sigma_n^2+\ln R(0)^2\sigma_g^2\sum_{n=1}^{\mathbb{N}}w_{in}\theta_n^2+w_{ig}\ln R(0)^2\theta_n\theta_g\sigma_g^2$$
$$-0.5\ln R(0)^2\left(\sigma_n^2+\theta_n^2\sigma_g^2\right)=\mathbb{E}_1\ln R_n(1)\ln W_{i1}(1)+\ln R(0)^2\tau_{in}w_{in}\sigma_n^2$$
$$+\ln R(0)^2\sigma_g^2\theta_n\left(\sum_{n=1}^{\mathbb{N}}w_{in}\theta_n+w_{ig}\theta_g\right)-0.5\ln R(0)^2\left(\sigma_n^2+\theta_n^2\sigma_g^2\right).$$

Similarly for the second term:

$$\begin{aligned}\mathbb{E}_1\left(\ln R_{g2}(2)-\ln R_{f2}(2)\right)&=\mathbb{E}_1\ln R_g(1)\ln W_{i1}(1)+\textstyle\sum_{n=1}^{\mathbb{N}}w_{in}\mathbb{E}_1\ln R_g(1)\ln R_n(1)\\&+w_{ig}\mathbb{E}_1\ln R_g(1)\ln R_g(1)+w_{if}\mathbb{E}_1\ln R_g(1)\ln R_f(1)-0.5\ln R(0)^2\theta_g^2\sigma_g^2\\&=\mathbb{E}_1\ln R_g(1)\ln W_{i1}(1)+\ln R(0)^2\theta_g\sigma_g^2\left(\textstyle\sum_{n=1}^{\mathbb{N}}w_{in}\theta_n+w_{ig}\theta_g\right)-0.5\ln R(0)^2\theta_g^2\sigma_g^2.\end{aligned}$$

Multiplying the latter by θ_n/θ_g and subtracting from the first term, one arrives at

$$\begin{aligned}\mathbb{E}_1\left(\ln R_n(2)-\ln R_f(2)\right)-\tfrac{\theta_n}{\theta_g}\mathbb{E}_1\left(\ln R_{g2}(2)-\ln R_{f2}(2)\right)&=\mathbb{E}_1\left(\ln R_n(1)-\tfrac{\theta_n}{\theta_g}\ln R_g(1)\right)\ln W_{i1}(1)\\&+\ln R(0)^2\tau_{in}w_{in}\sigma_n^2-0.5\ln R(0)^2\left(\sigma_n^2+\theta_n^2\sigma_g^2\right)+0.5\ln R(0)^2\theta_n\theta_g\sigma_g^2.\end{aligned}$$

Okawa and van Wincoop (2012) treat wealth as the zero order variable – in that case, we can bring expectation operation inside the first term (as there are no stochastic terms), and recall that $\mathbb{E}_1\ln R_{n2}(1)=\mathbb{E}_1\ln R_{g2}(1)=\mathbb{E}_1\ln R_{f2}(1)=0$, therefore,

$$\begin{aligned}w_{in}&=\frac{\mathbb{E}_1\left(\ln R_n(2)-\ln R_f(2)\right)-\frac{\theta_n}{\theta_g}\mathbb{E}_1\left(\ln R_{g2}(2)-\ln R_{f2}(2)\right)}{\ln R(0)^2\tau_{in}\sigma_n^2}\\&+\frac{0.5\ln R(0)^2\sigma_n^2+0.5\ln R(0)^2\theta_n^2\sigma_g^2-0.5\ln R(0)^2\theta_n\theta_g\sigma_g^2}{\ln R(0)^2\tau_{in}\sigma_n^2},\end{aligned}$$

where the second term goes to zero if $\frac{\sigma_n^2}{\sigma_g^2}=\theta_n(\theta_g-\theta_n)$. This equality implies that the weight used in the optimal share is $\frac{\theta_n}{\theta_g}=\left(\theta_g-\frac{\sigma_n^2}{\sigma_g^2}\frac{1}{\theta_n}\right)\frac{1}{\theta_g}$.[5] Hence, we arrive at $w_{in}=\frac{\mathbb{E}_1\left(\ln R_n(2)-\ln R_f(2)\right)-\frac{\theta_n}{\theta_g}\mathbb{E}_1\left(\ln R_{g2}(2)-\ln R_{f2}(2)\right)}{\ln R(0)^2\tau_{in}\sigma_n^2}$, as reported in the main text.

9 Econometric Methodology

9.1 Estimating Relations

We first need to connect decisions made by firms and agents into one estimable setting. Ignoring constant terms and time subscripts, the main relations are given by

$$\begin{cases}M_{in}^{FDI}=v_n^{1-\sigma}V_{in}^{FDI}\frac{E_iE_n}{[P_n\Pi_i]^{1-\sigma}}, & M_{in}^{I}=\gamma_{in}(\tau_{in})\frac{E_iE_n}{W\mathit{\Pi}_n\mathcal{P}_i}\delta_{in},\\ \frac{M_{in}^{I}}{M_{in}^{FDI}}=\frac{\gamma_{in}(\tau_{in})}{v_n^{1-\sigma}}\frac{[P_n\Pi_i]^{1-\sigma}}{W\mathit{\Pi}_n\mathcal{P}_i}\frac{\delta_{in}}{V_{in}^{FDI}}, & M_{in}^{I}+M_{in}^{FDI}=\left[\frac{v_n^{1-\sigma}V_{in}^{FDI}}{\{P_n\Pi_i\}^{1-\sigma}}+\frac{\gamma_{in}(\tau_{in})\delta_{in}}{W\mathit{\Pi}_n\mathcal{P}_i}\right]E_nE_i.\end{cases}$$

9.2 Probit Specification

We start with the bilateral FDI value. The log-linearized version of Z_{in} is

$$\begin{aligned}\mathsf{z}_{in}&=\gamma_0-\sigma\ln v_n+X_n-\phi_i-\phi_n-\kappa\phi_{in}+v_{in}\\&=\gamma_0-\phi_i-\phi_n-\sigma\ln v_n-\kappa\phi_{in}+X_n+\eta_{in},\end{aligned}\qquad(4)$$

where $f_{in}^{FDI}=\exp(\phi_i+\phi_n+\kappa\phi_{in}-v_{in})$ is a stochastic fixed investment cost, $\eta_{in}\equiv u_{in}+v_{in}\sim N\left(0,\,\sigma_u^2+\sigma_v^2\right)$ is i.i.d. (yet correlated with the error term u_{in} in the gravity equation at intensive margin), ϕ_i is an investor's fixed effect, whereas ϕ_n is a destination country's fixed effect, and X_n shall capture global effects, summarized in P_n and R_n, which are functions of the same underlying variables that describe all the partners of the destination country (i.e., P_n is the generalized mean of all prices that the destination economy has business with, whereas R_n includes the revenues from all the n's foreign markets; hence, X_n is captured by the cross-sectional averages in the Mundlak Chamberlain specifications). These fixed effects are expected to control for the country-specific and the time-invariant part of multilateral

[5] This, in turn, demonstrates that $\theta_g=1+\sigma_n^2/\sigma_g^2$, provided we imposed $\theta_n=1$.

resistance terms. Note that $\mathsf{z}_{in} > 0$ when i invests to n, and $\mathsf{z}_{in} = 0$ when it does not. Scaling z_{in} by the standard deviation σ_{u+v}, we obtain the probability ρ_{in} that i decides invest to n, conditional on the observed variables,

$$\begin{aligned}\rho_{in} &= \Pr\left(\mathsf{z}^{\star}_{in} > 0 \mid \text{observed variables}\right) \\ &= \Phi\left(\gamma^{\star}_{0} + \phi^{\star}_{i} + \phi^{\star}_{n} - \sigma^{\star} \ln \upsilon_{n} - \kappa^{\star}\phi_{in} + X^{\star}_{n}\right),\end{aligned} \tag{5}$$

where the starred variables refer to original ones divided by σ_{u+v} and $\phi^{\star}_{i}$, $\phi^{\star}_{n}$ denote the fixed effects.[6] Normalization is justified since the coefficients of a probit can only be estimated up to a scale. Moreover, the probit equation can be used to derive consistent estimates of W_{in} (which is crucial to avoid an omitted variable problem that biases the estimates). Further, note that inverse Normal maps estimated probability to the latent variable z_{in}, $\hat{\mathsf{z}}^{\star}_{in} = \Phi^{-1}\left(\hat{\rho}_{in}\right)$. Note that η_{in}/σ_{u+v} has a unit normal distribution, therefore, in principle, a consistent estimate is obtainable from the inverse Mills ratio (selection hazard). We can generate it from the estimation of a probit model where the error term follows a standard normal distribution. Hence, the approximate estimate of $\mathbb{E}\left(\eta^{\star}_{in} \mid \cdot,\, z^{\star}_{in} > 0\right)$ is simply $\phi\left(\hat{z}^{\star}_{in}\right)/\Phi\left(\hat{z}^{\star}_{in}\right)$. The estimate $\ln\left\{\exp\left(\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}\left(\hat{z}^{\star}_{in} + \widehat{\mathbb{E}\left(\eta^{\star}_{in} \mid \cdot,\, z^{\star}_{in} > 0\right)}\right)\right) - 1\right\}$[7] helps controlling for the potential correlation between the investment frictions and the average productivity of firms in country i which decide to invest to n (note that investment barriers affect cutoff productivity level). However, this method works only under homoskedasticity, a requirement rarely met in practice with bilateral flows data. As a result, Silva and Tenreyro (2009) claim that conditional expectation of $\frac{\eta_{in}}{\sigma_{u+v}} \equiv \frac{u_{in}+v_{in}}{\sigma_{u+v}}$, $\mathbb{E}\left(\eta^{\star}_{in} \mid \cdot,\, z^{\star}_{in} > 0\right)$ is inappropriately corrected.[8] As is obvious from (6), $\eta^{\star}_{in}$ is replaced by the conditional expectation and parameter $\beta_{u\eta} = \rho_{u,\eta}\frac{\sigma_u}{\sigma_v}$. The reason lies in Jensen's inequality, i.e., for any non-linear function $f(\cdot)$, $f\left(\mathbb{E}\left(\eta^{\star}_{in} \mid \cdot,\, z^{\star}_{in} > 0\right)\right) = f\left(\phi\left(z^{\star}_{in}\right)/\Phi\left(z^{\star}_{in}\right)\right) \neq \mathbb{E}\left(f\left(\eta^{\star}_{in}\right) \mid \cdot,\, z^{\star}_{in} > 0\right)$. Therefore, $f\left(\phi\left(\hat{z}^{\star}_{in}\right)/\Phi\left(\hat{z}^{\star}_{in}\right)\right)$ is not a consistent estimator. However, the approximation is justified if deviation form the conditional mean, $\eta^{\star}_{in} - \phi\left(z^{\star}_{in}\right)/\Phi\left(z^{\star}_{in}\right)$, is linear in the random term because it would be absorbed by the error of the equation. Notably, this approximation is likely to be reasonably accurate in many empirical studies.

Fortunately, it is possible to test for heteroskedasticity in the probit specification. Note that $\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}$ is proportional to the standard deviation of the error in (5). Hence, adding additional regressors $z^{\star 2}_{in}$ and $z^{\star 3}_{in}$ to the probit specification and checking for the joint significance would resemble a special case of White's test for heteroskedasticity. The test is analogous to a two-degrees-of-freedom RESET test (Ramsey, 1969), is also particularly interesting in that it can be interpreted as a normality test (see Newey, 1985). Therefore, this simple test provides a direct check for the validity of the main distributional assumptions required for consistent estimation of the model of interest. RESET test (based on a linear index) for the probit model should first be applied for the the validity of the results obtained with the two-stage estimators.[9] We skip full numerical results in the interest of space.

Substituting estimates into the FDI gravity equation leads to (here $\boldsymbol{X}$ entails all covariates and shall

---

[6] In empirical trade gravity literature, fixed effects are interpreted as controlling for multilateral resistance terms (which embody trade costs across trading partners). Whereas in the regional science literature, the fixed effects are treated as the "repulsive" and "attractive" forces of the exporter and the importer, see, for example, Roy (2004). In this application, fixed effects control for the system-wide barriers of investment as embodied in the price indices (refer to the main text for the expressions).

[7] The selection equation has been derived from a firm-level decision, and it therefore does not contain the unobserved and endogenous variable $W_{in} \equiv \max\left\{Z_{in}^{\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}} - 1,\, 0\right\}$ that is related to the fraction of multinationals (direct investors). The estimate allows mapping $\mathbb{E}\left(w^{\star}_{in} \mid \cdot,\, z^{\star}_{in} > 0\right)$ into the reported expression.

[8] Notably, $\eta^{\star}_{in}$ enters the model non-linearly that further complicates analysis - an estimator that is robust to incidental distributional assumptions is non-existent. More importantly, since $\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}$ depends on σ_u and σ_v, if either of the random components of the model is heteroskedastic, this makes the whole term a function of the regressors, which will then enter the model in a much more complex form.

[9] Presence of heteroskedasticity highly complicates the identification of the effects of the covariates in the intensive and extensive margins as pointed out by Silva and Tenreyro (2009).

not be confused with X_n above)

$$m_{in} = \boldsymbol{\beta}'\boldsymbol{X} + \ln\left\{\exp\left(\frac{\sigma_{u+v}\left(k-\sigma+1\right)}{\sigma-1}\left(\hat{z}^{\star}_{in} + \mathbb{E}\left(\eta^{\star}_{in}\,\widehat{\mid \cdot,\, z^{\star}_{in}} > 0\right)\right)\right) - 1\right\} + \beta_{u\eta}\mathbb{E}\left(\eta^{\star}_{in}\,\widehat{\mid \cdot,\, z^{\star}_{in}} > 0\right) + e_{in}, \tag{6}$$

where $\mathbb{E}\left(u_{in} \mid \cdot,\, z^{\star}_{in} > 0\right) = \beta_{u\eta}\mathbb{E}\left(\eta^{\star}_{in} \mid \cdot,\, z^{\star}_{in} > 0\right)$, $\beta_{u\eta} = \rho_{u,\eta}\frac{\sigma_u}{\sigma_v}$ and e_{in} is an *iid* error term $\mathbb{E}\left(e_{in} \mid \cdot,\, z^{\star}_{in} > 0\right) = 0$. Note that the term $\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}$ contains a relationship between Pareto shape parameter k which is directly related to firm heterogeneity and as such affects the variance of the error term for different values of k.[10] It is also related to the measure of the extensive margin, which is associated with the elasticity of substitution σ.[11] Further, when $\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}$ increases, the following approximation improves:

$$\ln\left\{\exp\left(\frac{\sigma_{u+v}\left(k-\sigma+1\right)}{\sigma-1}\left(\hat{z}^{\star}_{in} + \mathbb{E}\left(\eta^{\star}_{in}\,\widehat{\mid \cdot,\, z^{\star}_{in}} > 0\right)\right)\right) - 1\right\} \approx \frac{\sigma_{u+v}\left(k-\sigma+1\right)}{\sigma-1}\left(\hat{z}^{\star}_{in} + \mathbb{E}\left(\eta^{\star}_{in}\,\widehat{\mid \cdot,\, z^{\star}_{in}} > 0\right)\right).$$

Hence, in this case the terms enter (6) approximately linearly. Helpman, Melitz, and Rubinstein (2008) estimate (6) using nonlinear least squares as parameters enter nonlinearly in the expression for $w^{\star}_{in}$. The use of $\mathbb{E}\left(\eta^{\star}_{in} \mid \cdot,\, z^{\star}_{in} > 0\right)$ to control for $\mathbb{E}\left(u_{in} \mid \cdot,\, z^{\star}_{in} > 0\right)$ is the Heckman (1979) correction for sample selection. Notice, however, that the unobserved firm-level heterogeneity is corrected by the additional control $\hat{z}^{\star}_{in}$. At first, (4) is estimated, then proceeded to (6). To avoid reliance on the normality assumption for the unobserved capital frictions, valid excluded variables for the second stage are needed. Observe that $\hat{z}^{\star}_{in}$ enters both (5) and (6). The exogenous source of variation that enters only (5) and helps breaking collinearity problem is ϕ_{in}. Economically, this exclusion restriction concerns fixed investment costs which have bearing on variable transaction costs. In other words, we need variables which correlate with the z_{in} but not with the residual of the second stage equation. Further, observe that $z^{\star}_{in}$ includes only destination fixed costs which can be fruitfully exploited to estimate FDI gravity.

Note that the statistical implications of the log transformation of multiplicative models was well understood at least from Teeken and Koerts (1972). Recently, Jia and Rathi (2008) proposed an anti-log transformation which reduces the biases of exponentiation in the models with heteroskedastic errors.[12]

9.3 Tobit Specification

We deviate from the standard approach, and consider a tobit specification. It is especially suited for international investment when we allow for investor's decisions at both margins to be correlated. In other words, we introduce a possibility for the intensive and extensive margins, after controlling for covariates, to be driven by common factors.

[10] In particular, a greater dispersion of productivity (a lower k) decreases the concentration of firms around the cutoff. This implies a smaller decrease in the mass of operating firms for a given increase in the productivity cutoff.

[11] The change in W_{in}, as defined in Footnote 7, with respect to k is always positive, though it is not straightforward to establish the speed of change, $\frac{\partial W_{in}}{\partial k} = \frac{\sigma^2_{u+v}(k-\sigma+1)}{(\sigma-1)^2} Z_{in}^{\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}} \ln Z_{in} > 0$ since $Z_{in} > 1$ and $\frac{\partial^2 W_{in}}{\partial k^2} = \frac{\sigma^2_{u+v}}{(\sigma-1)^2} Z_{in}^{\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}} \ln Z_{in}\left(1 + \frac{\sigma_{u+v}(k-\sigma+1)}{(\sigma-1)}\ln Z_{in}\right)$ where cutoff productivities were assumed independent of k. The second derivative is negative whenever $\sigma_{u+v} < -\frac{\sigma-1}{k-\sigma+1}\left(\ln Z_{in}\right)^{-1}$. Only in this case larger k (or lower firm heterogeneity, see Chaney, 2008) induces an increase (at a diminishing rate) of extensive margin.

[12] With the error term normality, $\mathsf{m}_{in} = \mu_n + \boldsymbol{\beta}'\boldsymbol{x}_{in} + \varepsilon_{in}$ where $\mathsf{m}_{in} \equiv \ln M_{in}$ and $\varepsilon_{in} \sim N\left(0,\, \sigma^2_n\right)$, we have $\mathbb{E}\left(\mathsf{m}_{in}\right) = \exp\left(\tilde{\mu_n} + \tilde{\boldsymbol{\beta}}'\boldsymbol{x}_{in} + \frac{\tilde{\sigma^2_n}}{2}\right)$. Note that the variables with tilde must be consistent for $\mathbb{E}\left(\mathsf{m}_{in}\right)$ to be also consistent. However, this expression over-estimates the expected value because of convexity of exponential function. However, ignoring correlations between $\tilde{\mu_n}$, $\tilde{\beta}_n$ and $\tilde{\sigma^2_n}$, one can write: $\mathbb{E}\left(\mathsf{m}_{in}\right) = \exp\left(\tilde{\mu_n}\right)\prod_{n=0}^{I}\exp\tilde{\beta}_n\boldsymbol{x}_{in}\exp\left(\frac{\tilde{\sigma^2_n}}{2}\right)$ where $\mathbb{E}\exp\left(\tilde{\mu_n}\right) = \exp\left(\mu_n + \frac{\sigma\tilde{\mu}_n}{2}\right)$ and $\mathbb{E}\exp\left(\tilde{\beta}_n\boldsymbol{x}_{in}\right) = \exp\left(\beta_n\boldsymbol{x}_{in} + \frac{\left(x_{in}\tilde{\sigma}\beta_n\right)^2}{2}\right)$, therefore, $\widetilde{\mathbb{E}\left(\mathsf{m}_{in}\right)} = \exp\left(\tilde{\mu_n} - \frac{\tilde{\sigma}^2_{\tilde{\mu_n}}}{2}\right)\prod_{n=0}^{I}\exp\left(\tilde{\beta}_n\boldsymbol{x}_{in} - \frac{\left(x_{in}\tilde{\sigma}\beta_n\right)^2}{2}\right)\exp\left(\frac{\tilde{\sigma^2_n}}{2}\right)$, where the adjustment for $\exp\left(\tilde{\sigma^2_n}/2\right)$ is omitted for simplicity.

9.3.1 Two-Part Model

However, the standard tobit does not suit as it restricts the selection mechanism to be from the same model as that generating the outcome variable. Moreover, failure of homoskedasticity and Normality in the tobit model has serious consequences. The part models are more flexible and less prone to deviations from these assumptions.[13] The statistical problem, acknowledged by Felbermayr and Kohler (2006) in their corner-solutions version of the gravity equation, is that the conditional mean of actual flow m_{in} cannot be linear in explanatory variables, because there is a positive probability mass at $\mathsf{m}_{in} = 0$. One of the possible solutions - nonlinear least squares (NLS) - are imperfect because of potential heteroskedasticity (corner outcomes) in which case NLS are inefficient. More importantly, the coefficients obtained by NLS estimation of a model for $\mathbb{E}(\ln M_{in} \mid \cdot)$ are hard to give an adequate interpretation. NLS do not allow us to empirically identify the intensive and extensive margins of world investment patterns.

Let's first start with the observation that the observed capital flows can be treated as continuously distributed, non-negative latent variable. Given the fixed costs of accessing foreign economies, the decision of investing will be determined by the latent variables Z_{in} or δ_{in}. Let's first start with the assumption employed by Felbermayr and Kohler (2006) that intensive and extensive margins are independent conditional on all explanatory variables $\boldsymbol{X}_{in}$, denoted by $D(M^{\star}_{in} \mid Z_{in}, \boldsymbol{X}_{in}) = D(M^{\star}_{in} \mid \boldsymbol{X}_{in})$ where $D(\cdot \mid \cdot)$ is the conditional distribution, we end up with the two-part or hurdle model, see Wooldridge (2010). A truncated normal hurdle model by Cragg (1971) with the latent variable (positive capital flows in our analysis) is assumed to have a truncated normal distribution. Since we are mainly interested in recovering $\mathbb{E}(M_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0)$, we may not make distributional assumptions. One method is the smearing estimate by Duan (1983). However, a direct method would be to estimate $\mathbb{E}(M_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0)$ as

$$\mathbb{E}(M_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0) = \exp(\boldsymbol{X}_{in}\beta),$$

and this contains $M^{\star}_{in} = \exp(\boldsymbol{X}_{in}\beta + u)$. A quasi-MLE is (Fisher) consistent given the conditional mean is correctly specified. Note that capital flows need not correspond to the chosen density, what only needed is the same range in the conditional mean as allowed in the chosen density of linear exponential family (LEF). For every QMLE in LEF there is an asymptotically (or $\sqrt{N}$) equivalent weighted nonlinear least squares (WNLS) estimator if the conditional mean is well specified despite the conditional variance. Hence, we first estimate a probit $\Pr(M_{in} > 0 \mid \boldsymbol{X}_{in}) = \Phi(\boldsymbol{X}_{in}\boldsymbol{\gamma})$, then obtain QMLE estimates of $\mathbb{E}(M_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0) = \exp(\boldsymbol{X}_{in}\beta)$. Finally, we can estimate $\mathbb{E}(M_{in} \mid \boldsymbol{X}_{in}) = \Phi(\boldsymbol{X}_{in}\boldsymbol{\gamma})\exp(\boldsymbol{X}_{in}\beta)$.

9.3.2 Correlation Between The Margins

As mentioned, the assumption of independent mechanisms generating intensive and extensive margins seems to be theoretically unjustified. We employ the type II Tobit model which allows for common unobserved factors thereby extending type I Tobit model. Let $M^{\star}_{in} = \exp(\boldsymbol{X}_{in}\beta + u)$ and $\{\Pr(M_{in} > 0 \mid \boldsymbol{X}_{in})\} = 1(\boldsymbol{X}_{in}\boldsymbol{\gamma} + v > 0)$, and allow u and v be correlated. That is, let correlation be $\rho = Cov(u, v)/\sigma$ where $\sigma = \sqrt{Var(u)}$. Moreover, let $D(\ln(M^{\star}_{in})) = N(\boldsymbol{X}_{in}\beta, \sigma^2)$. To find the density $f(M_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0)$, we, as in Wooldridge (2010), use the change-of-variables formula, leading to

$$f(M_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0) = g(\mathsf{m}_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0)/M_{in},$$

[13] Flexibility comes from the fact that tobit is just a probit model combined with a truncated regression model, with the coefficient vectors in the two models restricted to be proportional to each other. The second stage can be modeled using $\mathbb{E}(\ln M_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0)$ where the crucial assumption being linearity in $\boldsymbol{X}_{in}$ (alternatively, full maximum likelihood can be used for efficiency reasons but this requires that errors are jointly normally distributed). Given the interest concentrates on the conditional expectation only, the log of capital can be estimated as exponential using quasi-MLE without further distributional assumptions. Moreover, as the theoretical motivation purports, the set of variables affecting decision to invest and the volume are not identical. This makes results more robust as identification strategy does not rest entirely on distributional assumptions and the nonlinearity of the inverse of Mills ratio.

where $g(\cdot \mid \boldsymbol{X}_{in},\, M_{in} > 0)$ is the conditional density of $M_{in}^{\star}$. By Bayesian rule,

$$\Pr(M_{in} > 0 \mid \boldsymbol{X}_{in}) \;\; g(\mathsf{m}_{in} \mid \boldsymbol{X}_{in},\, M_{in} > 0) = \Pr(M_{in} > 0 \mid \mathsf{m}_{in},\, \boldsymbol{X}_{in})\, h(\mathsf{m}_{in} \mid \boldsymbol{X}_{in}),$$

where $h(\mathsf{m}_{in} \mid \boldsymbol{X}_{in})$ is the conditional density of $\ln(M_{in})$. Let $\mathbf{1}(\boldsymbol{X}_{in}\boldsymbol{\gamma} + v > 0) = \mathbf{1}(\boldsymbol{X}_{in}\boldsymbol{\gamma} + (\rho/\sigma)\, u + e > 0)$ where $v = (\rho/\sigma)\, u + e$, $u = \mathsf{m}_{in} - \boldsymbol{X}_{in}\beta$ and $e \mid \boldsymbol{X}_{in},\, u \sim N(0,\, 1-\rho^2)$. This leads to

$$\begin{aligned} \Pr(M_{in} > 0 \mid \mathsf{m}_{in},\, \boldsymbol{X}_{in}) &= \tfrac{\Pr(M_{in}>0 \mid \boldsymbol{X}_{in}) g(\mathsf{m}_{in} \mid \boldsymbol{X}_{in},\, M_{in}>0)}{h(\mathsf{m}_{in} \mid \boldsymbol{X}_{in})} \\ &= \Phi\left([\boldsymbol{X}_{in}\boldsymbol{\gamma} + (\rho/\sigma)(\mathsf{m}_{in} - \boldsymbol{X}_{in}\beta)]\, (1-\rho^2)^{-\frac{1}{2}} \right). \end{aligned}$$

Recalling that $h(\mathsf{m}_{in} \mid \boldsymbol{X}_{in}) \sim N(\boldsymbol{X}_{in}\beta,\, \sigma^2)$ leads to

$$\begin{aligned} &f(M_{in} \mid \boldsymbol{X}_{in}) \\ &= \Phi\left([\boldsymbol{X}_{in}\boldsymbol{\gamma} + (\rho/\sigma)(\mathsf{m}_{in} - \boldsymbol{X}_{in}\beta)]\, (1-\rho^2)^{-\frac{1}{2}} \right) \phi\left((\mathsf{m}_{in} - \boldsymbol{X}_{in}\beta) / \sigma M_{in} \right). \end{aligned}$$

Combining with the density for $M_{in} = 0$, we obtain the log-likelihood

$$\begin{aligned} &l_i(\theta) = 1[M_{in} = 0] \ln[1 - \Phi(\boldsymbol{X}_{in}\boldsymbol{\gamma})] + 1[M_{in} > 0] \\ &\times \left\{ \ln\left[\Phi\left([\boldsymbol{X}_{in}\boldsymbol{\gamma} + (\rho/\sigma)(\mathsf{m}_{in} - \boldsymbol{X}_{in}\beta)]\, (1-\rho^2)^{-\frac{1}{2}} \right) \right] + \ln[\phi((\mathsf{m}_{in} - \boldsymbol{X}_{in}\beta)/\sigma)] - \ln\sigma - \mathsf{m}_{in} \right\}. \end{aligned}$$

Note that $\rho = 0$ yields a standard lognormal hurdle model. The challenge with this model, in which $(u,\, v)$ is independent of $\boldsymbol{X}_{in}$ and jointly normal, is the identification. Given the set of explanatory variables in $M_{in}^{\star} = \exp(\boldsymbol{X}_{in}\beta + u)$ are the same as in $\{\Pr(M_{in} > 0 \mid \boldsymbol{X}_{in})\} = 1(\boldsymbol{X}_{in}\boldsymbol{\gamma} + v > 0)$. Note that the latter term is a function of $(\boldsymbol{X}_{in},\, v)$, the law of iterated expectations yields

$$\begin{aligned} \mathbb{E}(\mathsf{m}_{in} \mid \boldsymbol{X}_{in},\, v) = \boldsymbol{X}_{in}\beta + \mathbb{E}(u \mid \boldsymbol{X}_{in},\, v) = \boldsymbol{X}_{in}\beta + \mathbb{E}(u \mid v) = \boldsymbol{X}_{in}\beta + \rho\sigma v, \\ \mathbb{E}(v \mid \boldsymbol{X}_{in},\, 1(\boldsymbol{X}_{in}\boldsymbol{\gamma} + v > 0) = 1) = \lambda(\boldsymbol{X}_{in}\boldsymbol{\gamma}) \equiv \tfrac{\phi(\boldsymbol{X}_{in}\boldsymbol{\gamma})}{\Phi(\boldsymbol{X}_{in}\boldsymbol{\gamma})}, \\ \mathbb{E}(\mathsf{m}_{in} \mid \boldsymbol{X}_{in},\, 1(\boldsymbol{X}_{in}\boldsymbol{\gamma} + v > 0) = 1) = \boldsymbol{X}_{in}\beta + \rho\sigma\lambda(\boldsymbol{X}_{in}\boldsymbol{\gamma}), \\ \mathbb{E}(\mathsf{m}_{in} \mid \boldsymbol{X}_{in},\, M_{in} > 0) = \boldsymbol{X}_{in}\beta + \rho\sigma\lambda(\boldsymbol{X}_{in}\boldsymbol{\gamma}). \end{aligned}$$

If $\boldsymbol{\gamma}$ can be consistently estimated, then the last equation can be nominally identified. However, the nonlinearity of $\lambda(\boldsymbol{X}_{in}\boldsymbol{\gamma})$ can act as a poor identification strategy, in particular over linear ranges. With unrestricted β, $\lambda(\boldsymbol{X}_{in}\boldsymbol{\gamma})$ is a function of $\boldsymbol{X}_{in}$, arbitrarily close to linear if we relax a probit model on $\Pr(M_{in} > 0 \mid \boldsymbol{X}_{in})$. If the identification is entirely based on $\mathbb{E}(M_{in} \mid \boldsymbol{X}_{in})$, the lognormal hurdle (note that $\rho\sigma$ is not separately identified by $\mathbb{E}(M_{in} \mid \boldsymbol{X}_{in})$ since $\boldsymbol{X}_{in}$ includes a constant) and the ET2T models with the same set of regressors yield no differences. Hence, ET2T is more convincing when the covariates determining the participation decision strictly contain those affecting the amount decision. Then the model is

$$M_{in} = 1[\boldsymbol{X}_{in}\boldsymbol{\gamma} + v \geq 0] \times \exp\left(\tilde{\boldsymbol{X}}_{in}\tilde{\beta} + u \right),$$

where $\tilde{\boldsymbol{X}}_{in} \subset \boldsymbol{X}_{in}$ and both include unity as the first element. Given at least one exclusion restriction, $\tilde{\beta}$ and $\rho\sigma$ are better identified because $\lambda(\boldsymbol{X}_{in}\boldsymbol{\gamma})$ is not an exact function of $\tilde{\boldsymbol{X}}_{in}$. Note that a linear rather than exponential type II Tobit model allows for negative outcomes on M_{in}.

10 Intensive and Extensive Margins

As we focus on the developing (low and middle income) countries, the evolution of capital flows along intensive and extensive margins might be considerably different. To track the importance of these two aspects, we compute two two margins for our country set.

To measure the two margins, let us define $\mathcal{N}_{t,\,h}$ to denote the number of capital links of vintage h that are active at time t. The vintage of a capital link is defined as the earliest time at which capital movements occur between a specific pair of two countries, based on their financial openness. A country is said to be financially open if it receives capital from at least one foreign country. Total capital flow can, therefore, be written as

$$M_t = \sum_{h=t_0}^{t} \bar{M}_{t,\,h} \mathcal{N}_{t,\,h},$$

where $\bar{M}_{t,\,h}$ is the average bilateral capital flow based on active links of vintage h. Further, let $\bar{M}_t$ be the average volume of capital flows at time t across all vintages, $\bar{M}_t = \sum_{h=t_0}^{t} \bar{M}_{t,\,h} / \sum_{h=t_0}^{t} \mathcal{N}_{t,\,h}$. Let's start with a change in capital flows on the intensive margin, which is defined as the movements in capital volumes over the existing relationships (links):

$$\triangle M_t^{\text{int}} = \sum_{h=t_0}^{t-1} \left(\bar{M}_{t,\,h} - \bar{M}_{t-1,\,h} \right) \mathcal{N}_{t-1,\,h} = \sum_{h=t_0}^{t-1} \triangle \bar{M}_{t,\,h} \mathcal{N}_{t-1,\,h},$$

whereas the changes in the number of active links are defined as

$$\triangle N_t = \sum_{h=t_0}^{t-1} \left(\mathcal{N}_{t,\,h} - \mathcal{N}_{t-1,\,h} \right) = \sum_{h=t_0}^{t-1} \triangle \mathcal{N}_{t,\,h}.$$

One can easily check that a change in capital flows is defined as the sum of changes on intensive and extensive margins, $\triangle M_t = \triangle M_t^{\text{int}} + \triangle \mathcal{N}_t \bar{M}_t$,[14]

$$\begin{aligned} \triangle M_t &\equiv \triangle \bar{M}_t \mathcal{N}_{t-1} + \triangle \mathcal{N}_t \bar{M}_t \\ &= \left(\bar{M}_t - \bar{M}_{t-1} \right) \mathcal{N}_{t-1} + \left(\mathcal{N}_t - \mathcal{N}_{t-1} \right) \bar{M}_t = \mathcal{N}_t \bar{M}_t - \bar{M}_{t-1} \mathcal{N}_{t-1}. \end{aligned}$$

Therefore, changes on the extensive margin are "weighted" by end-of-period average flow volumes, while changes on the intensive margin are "weighted" by beginning-of-period numbers of links. Our temporal dimension is short, therefore, results are not to be significantly affected by time weights. The graphical representation of the above results is to be found in the main text, refer to the Figure in the Section on *Recent Facts*.

[14] Notice that $\triangle M_t^{\text{int}} = \sum_{h=t_0}^{t-1} \triangle \bar{M}_{t,\,h} \mathcal{N}_{t-1,\,h} = \triangle \left(\sum_{h=t_0}^{t} \bar{M}_{t,\,h} / \sum_{h=t_0}^{t} \mathcal{N}_{t,\,h} \right) \sum_{h=t_0}^{t-1} \mathcal{N}_{t-1,\,h} = \triangle \bar{M}_t \mathcal{N}_{t-1}$.

References

Baltagi, B. H., P. Egger, and M. Pfaffermayr (2008): "Estimating regional trade agreement effects on FDI in an interdependent world," *Journal of Econometrics*, 145(1-2), 194–208.

Behar, A., and B. D. Nelson (2009): "Trade Flows, Multilateral Resistance and Firm Heterogeneity," Discussion paper, University of Oxford.

Bergstrand, J. H., and P. Egger (2007): "A knowledge-and-physical-capital model of international trade flows, foreign direct investment, and multinational enterprises," *Journal of International Economics*, 73(2), 278 – 308.

Chaney, T. (2008): "Distorted Gravity: the Intensive and Extensive Margins of International Trade," *American Economic Review*, 98(4), 1707–1721.

Cragg, J. G. (1971): "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods," *Econometrica*, 39(5), 829–844.

Devereux, M. B., and A. Sutherland (2011): "Country Portfolios In Open Economy Macro Models," *Journal of the European Economic Association*, 9(2), 337–369.

Duan, N. (1983): "Smearing Estimate: A Nonparametric Retransformation Method," *Journal of the American Statistical Association*, 78(383), pp. 605–610.

Eaton, J., and A. Tamura (1994): "Bilateralism and Regionalism in Japanese and U.S. Trade and Direct Foreign Investment Patterns," *Journal of the Japanese and International Economies*, 8(4), 478–510.

Felbermayr, G. J., and W. Kohler (2006): "Exploring the Intensive and Extensive Margins of World Trade," *Review of World Economics*, 142, 642–674, 10.1007/s10290-006-0087-3.

Hasebe, T., and W. P. Vijverberg (2012): "A Flexible Sample Selection Model: A GTL-Copula Approach," IZA Discussion Papers 7003, Institute for the Study of Labor (IZA).

Heckman, J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), 153–161.

Helpman, E., M. J. Melitz, and Y. Rubinstein (2008): "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 123(2), 441–487, 2007.

Jia, S., and S. Rathi (2008): "On Predicting Log-transformed Linear Models with Heteroscedasticity," *SAS Global Forum*.

Newey, W. K. (1985): "Maximum Likelihood Specification Testing and Conditional Moment Tests," *Econometrica*, 53(5), 1047–70.

Okawa, Y., and E. van Wincoop (2012): "Gravity in International Finance," *Journal of International Economics*, 87(2), 205–215.

Ramsey, J. B. (1969): "Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis," *Journal of the Royal Statistical Society. Series B (Methodological)*, 31(2), pp. 350–371.

Roy, J. R. (2004): *Spatial Interaction Modeling: A Regional Science Context*. Berlin: Springer-Verlag.

Santos-Silva, J., and S. Tenreyro (2007): "The log of gravity," *Review of Economics and Statistics*, 88, 641–658.

Silva, J. M. C. S., and S. Tenreyro (2009): "Trading Partners and Trading Volumes: Implementing the Helpman-Melitz-Rubinstein Model Empirically," CEP Discussion Papers dp0935, Centre for Economic Performance, LSE.

Stewart, M. (2006): "Maximum simulated likelihood estimation of random-effects dynamic probit models with autocorrelated errors," *Stata Journal*, 6(2), 256–272.

Stewart, M. B. (2007): "The interrelated dynamics of unemployment and low-wage employment," *Journal of Applied Econometrics*, 22(3), 511–531.

Teeken, R., and J. Koerts (1972): "Some Statistical Implications of the Log Transformation of Multiplicative Models," *Econometrica*, 40(5).

Tille, C., and E. van Wincoop (2010): "International capital flows," *Journal of International Economics*, 80(2), 157–175.

Wooldridge, J. M. (2010): *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, second edn.