Appendix For Online Publication

Financing Multinationals

Jingting Fan
Pennsylvania State University

Wenlan Luo
Tsinghua University

Contents

A Data and Empirics

This section provides supplementary information on data sources and robustness results. A.1 and A.2 introduce the data for bilateral and firm-level analysis, respectively. A.3 reports the regressions based on the bilateral data, which complements Figure 1 of the text. A.4 reports bilateral evidence based on the data from the U.S. BEA, which sheds light on the mechanisms. Finally, A.5 provides the robustness results for firm-level regressions.

A.1 Bilateral Data and Sample Countries

Bilateral FDI and MP. Our primary source of data for the cross-sectional analysis is Ramondo et al. (2015). The dataset includes bilateral FDI stock and MP (sales and the number of affiliates), averaged over 1996-2001, and various measures of bilateral distance between 59 major countries. Bilateral affiliate sales for some country pairs are missing and hence imputed based on FDI and the number of affiliates in their dataset. These imputed values are excluded from our reduced-form analysis.

We supplement the data from Ramondo et al. (2015) with additional information, described below. For consistency in timing, whenever possible, we average these supplementary variables over 1996-2001. In the circumstance that a variable has missing values for all five years, we use the last available value of the variable before 1996 or—if no such value exists—the earliest available value after 2001.

Financial development index and private credit. Both the financial development index and the total private credit are obtained from the World bank. The total private credit measures the amount of loans made by domestic financial institutions—including banks and others—to local private enterprises. The financial development index is constructed based on surveys of practitioners, such as lawyers, consultants, and government officials, that work in finance-related areas. It consists of two sub-indices. The first, the depth of credit information index, measures the accessibility and quality of a country's credit information for lending decisions, and can take a value between 0 (poor) to 6 (good). The second, the strength of legal rights index, measures a country's protection of investors and lenders through collateral and bankruptcy laws, and can take a value between 0 (poor) and 10 (good). In the baseline analysis, we use the logarithm of the sum of the two indices to measure the qualify of the financial institutions.

GDP and TFP. Country TFP and real GDP are obtained from the Penn World Table (PWT) 9.0. We use PPP deflated (output-side) GDP and TFP (cgdpo and ctfp, respectively).

Business profit tax rate. This measure is obtained from the World Bank. It is calculated as the average total tax payable by domestic firms as percent of their profits.

Low tax/tax haven dummy. We create a tax-haven indicator for countries that are frequently associated with tax shielding. According to Appendix Table A7 of Torslov et al. (2018), the top countries in the *level* of shifted profits in our sample are: Ireland, Singapore, Switzerland, and the Netherlands. These four countries also rank among the top in the *share* of shifted profits in total corporate profits.[1] We assign a value of 1 to these countries and a value of 0 to everyone else.

FDI Regulatory Restrictiveness Index. The FDI Regulatory Restrictiveness Index, constructed by the OECD, measures countries' stringency of regulations on FDI along four dimensions: limitations on equity holding, discriminatory policies on screen and approval process, restrictions on employment of foreigners as key personnel, and other operational restrictions. The index is available at both country and sector level. We use the country-level index.

Sample countries. Our analysis focuses on 36 major developed and developing countries. This sample is selected from the 59 countries in Ramondo et al. (2015) as follows. First, we drop countries that are outside the top 35 in *all* of the following: GDP, total outward FDI, total inward FDI. This restriction leaves us with 39 countries that either is large in themselves or plays a disproportionately important role in world FDI. We further exclude Uruguay, El Salvador, Croatia, and Bulgaria because the BEA does not provide balance sheet and external finance information of American MNE affiliates in these countries, which is needed for our calibration of μ_h in quantification. Although not reported here, all empirical results are robust if the full sample in Ramondo et al. (2015) is used.

Table A.1 summarizes the sources and descriptive statistics of these variables.

Table A.1: Descriptive Statistics for the Cross-Sectional Data

	Source	N	Mean	Std. Dev.	Min	Max
Log (GDP)	PWT	36	13.25	1.20	11.48	16.31
Log (TFP)	PWT	36	0.78	0.24	0.27	1.21
FDI restrictiveness index	OECD	34	0.82	0.17	0.37	0.99
Credit/GDP (%)	WB	36	79.38	47.19	12.81	213.13
Business tax (% of profit)	WB	36	15.75	8.04	1.20	28.80z
Financial development index	WB	36	10.81	3.21	3.00	16.00
log(Financial development index)	WB	36	2.32	0.37	1.10	2.77

A.2 Descriptions of the Firm-level Data

The main data source of our firm-level analysis is Orbis, which covers the period of 2001-2016. Our cleaning procedures largely follow Cravino and Levchenko (2017) and Fan (2017). We summarize these procedures below for the convenience of readers; see the appendix of Fan (2017) for more details.

Accounting variables. Accounting data are available at firm identifier-year level. Each firm identifier corresponds to a unique entity that might be owned by other firms or individuals. Some firms report financial data from multiple sources (e.g. the local registry, annual reports). In such cases, we use the data from the local registry, which is deemed more accurate. Some firms report not only unconsolidated accounting statements but also consolidated statements that summarize the operation of entire business group to which the firm belongs. We will use only the information from unconsolidated accounts. The variables we extract are turnover (sales), wage bill, and total assets. Following Ramondo et al. (2015), we measure multinational production using affiliate sales, but we will use wage bill for robustness analysis.

[1] Some other offshore tax havens, such as Panama, Bermuda, Bahamas, Luxembourg, also have high *shares* of shifted profits, but they are not in our sample.

Ownership linkage and FDI. We use the Orbis ownership database to link firm identifiers to their 'global ultimate owner' (GUO), the firm/individual/family that ultimately owns a firm, which we also label as the parent of a firm. The GUO of a firm is defined as the entity that satisfies both of the following criteria. First, it owns more than 50% control of the firm either directly or indirectly; when the ownership is indirect, all intermediate owners on the ownership chain must hold more than 50% of the 'downstream' owners. Second, this entity does not have another entity that owns more than 50% of its share—in this case, the latter entity should be the 'ultimate' owner. For firms that are not otherwise linked to a GUO, we assume that their GUO are themselves, which practically means they are all domestic firms and will not be included in regressions (see our sample criteria below). The coverage of the ownership database expanded gradually but some firms might be in and out of the database. For broad representation, we use the latest snapshot of each firm to construct the ownership links.

As noted in the main text, the majority of firms in the ownership database report only the identity, but not the exact position, of their GUOs. Furthermore, the available information on firms' balance sheets is not detailed enough for us to measure the intra-company loans from parents to affiliates. We will use firms' total assets and a panel dataset of bilateral FDI stock to construct a firm-level FDI proxy. We explain the construction of the bilateral FDI stock in Appendix C.1.

Sample firms and period. Our regressions will restrict to firms that belong to multinational firms. With this restriction, our estimates are identified from the comparison among affiliates from and operating in different countries, rather than the comparison between foreign affiliates and domestic firms. Our baseline analysis focuses on 2001-2012 because only for this period we are able to construct bilateral FDI, which is needed in the proxy for firm-level FDI. We note that for all specifications that do not require the firm-level FDI proxy, results are similar when the full 2001-2016 firm-level data are used.

A.3 Regression Evidence Using Bilateral Data

Quality of financial institutions and inward MP. Tables A.2 reports the regressions on the relationship between the quality of financial institutions and inward MNE activities. To account for zeros and to avoid the bias from log transformation in the presence of heteroskedasticity, all specifications are estimated using the Poisson Pseudo Maximum Likelihood estimator.

The first column shows that controlling for the overall size and productivity of a host, its financial development index is positively associated with inward FDI. The second column controls for home fixed effects and four measures of bilateral distance—geographic distance and indicators for whether the home and the host share a border, a common official language, or a colonial tie—to absorb the variation due to the heterogeneity in host geographic locations. The coefficient remains essentially the same. The attractiveness of a country to foreign investors likely also depends on its profit tax rate and policy restrictions on foreign firms. The third column further controls for the FDI restrictiveness index, the average profit tax rate, and the indicator of whether a host is viewed as a 'tax-haven' country. The coefficient is smaller, but remains sizable and statistically significant. The point estimate implies that a one-standard-deviation increase in the logarithm of financial development index is correlated with a 34 log point increase in inward FDI. In Column 4, we split the financial development index into its two subcomponents, the protection for creditors' legal rights and the depth of credit information. A priori, because MNE affiliates are likely large firms well covered by the press and analysts' reports, it appears

Table A.2: Host Finance and Inward MNE Activities

	(1)	(2)	(3)	(4)	(5)	(6)
		log (FDI stock)			log (MP)	
log (host financial development index)	1.486***	1.271***	0.910***		0.591***	
	(0.425)	(0.393)	(0.350)		(0.219)	
log (creditor legal rights)				0.591**		0.340**
				(0.293)		(0.159)
log (credit info depth)				-0.301		0.135
				(0.407)		(0.246)
log (FDI stock)					0.559***	0.550***
					(0.066)	(0.070)
log host GDP	0.517***	0.661***	0.794***	0.814***	0.418***	0.430***
	(0.062)	(0.067)	(0.064)	(0.057)	(0.081)	(0.085)
host TFP	3.063***	2.427***	1.431*	1.092	-0.256	-0.312
	(0.801)	(0.717)	(0.800)	(0.825)	(0.285)	(0.281)
Observations	1171	1170	1104	1038	450	437
R^2	0.128	0.797	0.865	0.865	0.960	0.961
Controls on host policy			yes	yes	yes	yes
Home FE		yes	yes	yes	yes	yes
Bilateral distance		yes	yes	yes	yes	yes

All specifications are in PPML. 'Bilateral distance' include geographic distance and indicators of whether the home and the host share a border, a common official language, or a colonial tie. 'Controls on host policy' include hosts' FDI restrictiveness index, average profit tax rate, and tax haven indicator. Standard errors (in parenthesis) are clustered by host countries. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

unlikely that the depth of credit information will have an impact on their finance. On the other hand, laws and the court systems that are more favorable to investors, captured in the index for creditors' legal rights, might help these affiliates secure external finance. The result from Column 4 confirms this prior: only the index for the protection of credits' rights has a statistically significant correlation with FDI stock. In Columns 5 and 6, we examine the relationship between host financial development and inward MP, *conditional on* bilateral FDI. We find that FDI has a large coefficient, and that conditional on it, the quality of financial institutions is correlated with inward MP.

To summarize, Table A.2 documents two main findings. First, FDI is highly correlated with MP, indicating the importance of parent finance for affiliate production. Second, controlling for FDI, the quality of host financial institutions, measured either as a combined index or as the protection of creditor's legal rights, is strongly and positively correlated with MP.

Quality of financial institutions and outward MP. Table A.3 reports the results on the relationship between the quality of financial institutions and outward MNE activities. The first three columns focus on the stock of outward FDI as the outcome variable, gradually adding controls on host fixed effects and bilateral distances, and other factors that might indirectly affect domestic firms' incentive of going overseas, such as its tax rate, its restrictiveness on FDI, and its status as a tax haven. These columns show a robust correlation between the quality of home financial institutions and outward FDI. Column 4 shows that it is the protection of creditors' legal rights, as opposed to the depth of information in the credit market, that explains the correlation. This is broadly in line with our finding in Table A.2.

Columns 5 and 6 further show that conditional on bilateral FDI, neither the overall home financial index nor its subcomponents have a significant relationship with MP. Thus, the correlation between the quality of home financial institutions and outward MP is entirely accounted for by the level of FDI.

Table A.3: Home Finance and Outward MNE Activities

	(1)	(2)	(3)	(4)	(5)	(6)
	log (FDI stock)				log (MP)	
log (home financial development index)	1.831***	1.645***	1.328***		0.288	
	(0.603)	(0.271)	(0.311)		(0.400)	
log (credit info depth)				0.169		0.179
				(0.478)		(0.586)
log (creditor legal rights)				0.780***		0.169
				(0.200)		(0.337)
log (FDI stock)					0.822***	0.824***
					(0.050)	(0.046)
log home GDP	0.586***	0.798***	0.828***	0.841***	0.321***	0.315***
	(0.086)	(0.080)	(0.083)	(0.076)	(0.089)	(0.092)
home TFP	4.256***	3.872***	3.466***	3.299***	-1.312**	-1.318**
	(0.730)	(0.641)	(0.691)	(0.703)	(0.523)	(0.564)
Observations	1171	1170	1170	1108	459	450
R^2	0.147	0.850	0.862	0.866	0.956	0.956
Controls on home policy			yes	yes	yes	yes
host FE		yes	yes	yes	yes	yes
Bilateral distance		yes	yes	yes	yes	yes

All specifications are in PPML. 'Bilateral distance' include geographic distance and indicators for whether the home and the host share a border, a common official language, or a colonial tie. 'Controls on home policy' include homes' FDI restrictiveness index, average profit tax rate, and tax haven indicator. Standard errors (in parenthesis) are clustered by home countries. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.4 Evidence from the BEA Public-Use Data

This section provides details on the BEA data, which were used to show the source of external finance of U.S. MNE overseas affiliates in Section 2.2. In addition, we also replicate the findings in Figure 1 using this dataset. This exercise complements the evidence in Figure 1 and Section A.3 because we can use total compensations to measure the activities of affiliates, alleviating the concern that sales might be mis-measured or manipulated.

Data descriptions. We rely on the public-use table produced by the BEA from the firm-level surveys it administers. From the data we assemble three tables. **The first table** includes the sources of external finance for U.S. overseas affiliates, aggregated to the host country-year level—this is the source of information for Figure 2 of the host. **The second table** includes the investment position of U.S. parents in their overseas affiliates (in either equity and intra-company loans), and the operational and balance sheet statistics of these affiliates, all aggregated to the host country-industry-year level. **The third table** contains the basic balance sheet and operational statistics of the foreign affiliates operating in the U.S., aggregated by year and the home country of these affiliates.

Sample country and period. We focus on the same 36 countries as in the rest of the paper. The BEA changed variable definitions and industry classifications a few times over the past decades. For consistency and constrained by the availability of different tables, we focus on 1999-2007 for the relationship between the quality of host financial institutions and inward MNE activities and on 1999-2006 for the relationship between the quality of home financial institutions and outward MNE activities.[2] Given that our benchmark measure of financial institution quality does not vary over this period, we perform a pooled regression across years.

Quality of financial institutions and inward MNE activities. Table A.4 reports the regressions on the

[2] In quantification, we will use a time series over 2001-2012 constructed from the BEA data to pin down the time series of μ_h. There, we make additional assumptions to extend the time series. See Appendix C.1 for details.

Table A.4: Host Finance and Inward MNE Activities: Evidence from the U.S. BEA data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Operation and Balance Sheet				External Finance	
Dependent Variable	FDI	Sales	Wage bill	Asset	Wage bill	Parent	Host Ctry
log (financial development index)	1.572***	0.462**	0.360**	1.186***	0.133	1.196***	0.603***
	(0.330)	(0.195)	(0.168)	(0.223)	(0.218)	(0.357)	(0.196)
log (parent investment position)		0.615***	0.570***	0.756***			
		(0.040)	(0.035)	(0.077)			
log (total affiliate asset)					0.644***		
					(0.050)		
log (parent external finance)							0.839***
							(0.105)
Observations	4094	2272	2895	2305	2581	352	183
R^2	0.774	0.902	0.908	0.952	0.949	0.880	0.980
Home country FE	yes	yes	yes	yes	yes	yes	yes
Host GDP and TFP	yes	yes	yes	yes	yes	yes	yes
Additional host chars.	yes	yes	yes	yes	yes	yes	yes
Bilateral distance measures	yes	yes	yes	yes	yes	yes	yes
Industry-year FE	yes	yes	yes	yes	yes		
Year FE						yes	yes

Note: See Table A.2 for the definition of bilateral distance and host characteristics. See the text in Section A.4 for discussions. Standard errors (clustered by host countries) in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

relationship between the quality of host financial institutions and the activities of U.S. multinationals.[3] Columns 1 through 5 are based on the second data table described above. In addition to the full set of controls in the baseline analysis (home country fixed effects, host country characteristics and bilateral distance measures), we also include the industry-year fixed effects. This ensures that our findings are not driven by host countries with better financial institutions systematically specializing in industries in which MNEs are more prevalent. Columns 6 and 7 are based on the second data table described, which has no industry dimension, so we control only for country and year fixed effects.

Columns 1 and 2 of Table A.4 convey the same message as in Table A.2. The dependent variable in Column 1 is log of the total investment position of U.S. parents in affiliates. This measure maps closely into outward FDI in Table A.2 and, reassuringly, the point estimate is not far from the one in Column 2 of Table A.2 either. Column 2 uses log sales as the dependent variable and includes log of parent position. Coefficients for both financial development and the log of parent positions are similar to those in Column 5 of Table A.2, although the two specifications are estimated on two different datasets.

One concern for using affiliate sales to measure MP is that sales can be manipulated by firms for profit shifting (Guvenen et al., 2019). Moreover, some of the sales might reflect the value of intermediate inputs from the parents and thus fail to accurately capture affiliate production. Using total compensation as the dependent variable gives a similar result, as reported in Column 3.

Columns 4 and 5 provide evidence for the mechanism. Specifically, they show that conditional on parent investment, the quality of host financial institutions is strongly correlated with affiliates' total assets. Moreover, once the total assets are controlled for, the quality of host financial institutions is no longer correlated with affiliate production. Together, these findings suggest that affiliates produce more in countries with higher quality of financial institutions because they scale up more in these countries,

[3] Antras et al. (2009) document causal evidence supportive of this relationship using firm-level data from the same source. The results reported in this section complements Antras et al. (2009) by showing the relationship between the quality of home financial institutions and outward MNE activities.

Table A.5: Home Finance and Outward MNE Activities: Evidence from the BEA data

	(1)	(2)	(3)	(4)
		Operation and Balance Sheet Info		
Dependent Variable	FDI	Sales	Wage bill	Assets
log (financial development index)	3.171^{***}	-0.169	0.197	0.189
	(1.012)	(0.298)	(0.368)	(0.433)
log (parent investment position)		0.773^{***}	0.914^{***}	1.093^{***}
		(0.071)	(0.071)	(0.110)
Observations	315	249	277	254
R^2	0.847	0.974	0.977	0.973
Host country FE	yes	yes	yes	yes
Home GDP and TFP	yes	yes	yes	yes
Additional home chars.	yes	yes	yes	yes
Bilateral distance measures	yes	yes	yes	yes
Year FE	yes	yes	yes	yes

Note: See Table A.3 for definition of bilateral distance and home characteristics. All specifications are estimated using PPML. See the text in Section A.4 for discussions. Standard errors (clustered by home countries) in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

not because the quality of financial institutions is correlated with unobserved factors that improve the productivity of foreign affiliates.

Columns 6 and 7 of Table A.4 examine how affiliates' source of external finance vary with the quality of host financial institutions. The dependent variables are the log of total external finance from the parent and the host country, respectively. Both external finance measures are flows, and both include equity and loans. We find that the quality of host financial institutions is positively correlated with external finance from the parent, and that conditional on it, external finance from the host. Since external finance measures here are flows, these results also rule out the concern that the correlation we find is driven by MNEs' accumulation of past retained earnings in their affiliates.

Quality of financial institutions and outward MNE activities. Table A.5 reports the relationship between the quality of home financial institutions and outward MNE activities. Because external finance information is unavailable for most countries in the public-use data, we focus on operation and balance sheet variables. Column 1 of Table A.5 shows that countries with better financial institutions invest more in their affiliates in the U.S. Column 2 shows that, conditional on the parent investment position (FDI), the quality of home financial institutions is *not* correlated with the sales of their affiliates. Column 3 confirms this result with wage bill as the proxy for affiliate operation. Column 4 shows that when the dependent variable is the size of affiliate balance sheet, the coefficient for home financial institution quality is also small. This finding is consistent with our central mechanism that the quality of home financial institutions affecting MP through the impact of parent investment on affiliate size.

A.5 Additional Results from Firm-Level Data

Robustness. The first eight columns of Table A.6 reports the results from the robustness exercises, following the specifications of Columns 3, 4 and 7 of Table 1 but using different measures/controls.

Columns 1, 2, and 6 interact the measures of home and host country credit conditions with an indicator for the post-financial crisis period (2008-2012). The coefficients for these interaction terms are small, suggesting that the patterns do not differ systematically before and after the crisis. Columns 3 and 7 change the proxy for FDI in firm i at year t. Instead of using the proportional rule, as described in the text, we control for flexible interaction terms between the affiliates' lagged capital stock and bilateral FDI

Table A.6: Firm-level Regressions: Robustness

	Home credit and affiliate activities					Host credit and affiliate activities			Interaction
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent var.	log(sales)			log(wage bill)		log(sales)		log(wage bill)	log(sales)
$\text{credit}_{o(i)t}$	0.097***	0.016	0.002	0.033**	-0.035				
	(0.023)	(0.030)	(0.022)	(0.017)	(0.026)				
$\text{credit}_{o(i)t} \times \mathbb{I}_{t \geq 2008}$	-0.004	-0.005							
	(0.004)	(0.005)							
ln(parent sales)	0.028***	0.026***	0.025***	0.020***	0.018***				
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)				
$\widehat{\text{FDI}}_{i,t}$		0.188***	flexible		0.167***	0.179***	flexible	0.158***	
		(0.008)			(0.006)	(0.006)		(0.009)	
$\text{credit}_{d(i)t}$						0.208***	0.269***	0.122***	
						(0.056)	(0.054)	(0.044)	
$\text{credit}_{d(i)t} \times \mathbb{I}_{t \geq 2008}$						0.014**			
						(0.007)			
$\text{credit}_{o(i)t} \times \text{credit}_{d(i)t}$									-0.007
									(.005)
Observations	550837	393579	364368	434662	321740	378750	348817	309065	395,730
R^2	0.876	0.899	0.908	0.909	0.928	0.917	0.922	0.944	0.869
Affiliate FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
Host-year FE	yes	yes	yes	yes	yes	-	-	-	yes
Home economic shocks	yes	yes	yes	yes	yes	-	-	-	-
Firm-year FE	-	-	-	-	-	yes	yes	yes	yes
Host economic shocks	-	-	-	-	-	yes	yes	yes	-

Standard errors (in parenthesis) are clustered two-way, by host-year and home-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

at time t (we control for up to the third order). Columns 4, 5, and 8 replicate columns 3, 4, and 7 of Table 1 using affiliate wage bill as the measure of production.

These results corroborate the findings from our baseline analysis: home credit shocks are correlated with affiliate sales, but the correlation vanishes once firm-level FDI is controlled for; on the other hand, host credit shocks are strongly correlated with affiliate sales even after firm-level FDI is controlled for.

Non-interaction between parent and affiliate shocks. Our model predicts that although affiliate sales depend on both home and host credit shocks, when estimated in a log specification, there is neither positive nor negative interaction between these two shocks. In Column 9, we regress log affiliate sales on the interaction term, controlling for affiliate, firm-year, and host-year fixed effects. Consistently with the model, we find a small and rather precisely estimated coefficient.

B Theory

We define the sequential competitive equilibrium in B.1 and derive the equations and propositions in B.2. Section B.3 describes the model of MP without FDI used for welfare comparison. Section B.4 microfounds the frictions faced by parents and affiliates. Section B.5 shows the isomorphism of our model to a model with differentiated varieties. Section B.6 shows that the analytical tractability of our model carries over to the more general CRRA utility family and to cases with firm-level hysteresis in international investments.

B.1 Formal Definition of Equilibrium

We add back the time subscript to be explicit that we focus on the sequential equilibrium to study the transition path with time-varying parameters. The aggregate state of the economy is the joint distribution of parent firms' net worth and productivity in each country, characterized by joint density functions $\{\Phi_{i,t}(z,a)\}_{i=1}^{N}$. Using the property that the policy functions for parent firms' financing, investing and fund allocation are linear in net worth (Lemmas 2 and 3), and that the policy functions for affiliates' financing, production and factor usage are linear in the funds from the parents (Lemma 1), it is sufficient to track the measure of total net worth held by parent firms with productivity z in each country i (i.e., the wealth density function $\phi_{i,t}(z)$ defined in Section 3.5).

Definition 1. *Given initial density functions $\{\Phi_{i,0}(z,a)\}_{i=1}^{N}$, a sequential competitive equilibrium is a sequence of wealth density functions $\{\phi_{i,t}(z)\}_{t=0}^{\infty}$, wage and interest rates $\{w_{i,t}, r_{i,t}^{b}\}_{t=0}^{\infty}$, affiliates' return and policy functions $\{R_{ih,t}(z), \hat{b}_{ih,t}^{F}(z), \hat{k}_{ih,t}(z), \hat{l}_{ih,t}(z), \hat{y}_{ih,t}(z)\}_{t=0}^{\infty}$, parents' value, policy, and return functions $\{v_{i,t}(z,\boldsymbol{\zeta},a), c_{i,t}(z,\boldsymbol{\zeta},a), a'_{i,t}(z,\boldsymbol{\zeta},a), \hat{b}_{i,t}^{H}(z,\boldsymbol{\zeta}), R_{i,t}^{a}(z,\boldsymbol{\zeta}), \hat{e}_{ih,t}(z), \mathbb{E}[R_{i,t}^{a}(z,\boldsymbol{\zeta})|z]\}$, such that*

1. *Affiliates' return and policy functions solve affiliates' financing and production problems, characterized by Lemma 1. Parents' value, policy and return functions solve parents' financing, investing and fund allocation problems, characterized by Lemma 2 and 3.*

2. *In each country h, period t, the labor market clears by country:*

$$\sum_{i=1}^{N}\int_{0}^{\infty}\hat{l}_{ih,t}(z)(1+\lambda_{i,t})\hat{e}_{ih,t}(z)\phi_{i,t}(z)dz = L_{h,t}, \forall h$$

The global bond market clears:[4]

$$\underbrace{\sum_{i=1}^{N}\int_{0}^{\infty}\left[1-\sum_{h=1}^{N}\hat{e}_{ih,t}(z)\right]\phi_{i,t}(z)dz}_{\text{Bond supply from idle parent firms}} = \underbrace{\sum_{i=1}^{N}\int_{0}^{\infty}\lambda_{i,t}\sum_{h=1}^{N}\hat{e}_{ih,t}(z)\phi_{i,t}(z)dz}_{\text{Bond demand from active parent firms}} + \underbrace{\sum_{i=1}^{N}\sum_{h=1}^{N}\int_{0}^{\infty}\hat{b}_{ih,t}^{F}(z)(1+\lambda_{i,t})\hat{e}_{ih,t}(z)\phi_{i,t}(z)dz}_{\text{Bond demand from affiliates}}, \quad \text{(B.1)}$$

and $r_{i,t}^{b} = r_{i',t}^{b}\ \forall i, i'$.[5]

[4]Recall that bonds are for trading capital. By applying $\hat{k}_{ih,t}(z) = 1 + \hat{b}_{ih,t}^{F}(z)$ characterized in Lemma 1, equation B.1 is equivalent to the global capital market clearing condition: $\sum_{h=1}^{N}\sum_{i=1}^{N}\int_{0}^{\infty}\hat{k}_{ih,t}(z)(1+\lambda_{i,t})\hat{e}_{ih,t}(z)\phi_{i,t}(z)dz = \sum_{i}\int_{0}^{\infty}\phi_{it}(z)dz$. This equation also implies that total supply of productive capital (the sum of net worth) equals the total use of productive capital in the world economy.

[5]If we assume segregated bond markets, then the bond markets clear by country with country-specific interest rates.

The goods market clears by country:[6]

$$\overbrace{\sum_{i=1}^{N}\int_0^{\infty}[\hat{y}_{ih,t}(z)+(1-\delta)\hat{k}_{ih,t}(z)](1+\lambda_{i,t})\hat{e}_{ih,t}(z)\phi_{i,t}(z)dz}^{\text{output and un-depreciated capital}} \tag{B.2}$$
$$= \underbrace{w_{h,t}L_{h,t}}_{\text{worker consumption}} + \underbrace{\sum_{i=1}^{N}\int_0^{\infty}R_{ih,t}(z)(1+\lambda_{i,t})\hat{e}_{ih,t}(z)\phi_{i,t}(z)dz}_{\text{repatriated returns}} + \underbrace{(1+r^b_{h,t})\sum_{i=1}^{N}\int_0^{\infty}\hat{b}^F_{ih,t}(z)(1+\lambda_{i,t})\hat{e}_{ih,t}(z)\phi_{i,t}(z)dz}_{\text{interests paid to local creditors}}$$

Note that nominal values (e.g., wages and returns) enter the condition because the final consumption good is the numeraire. That is, these prices are in units of the consumption good. We assume that un-depreciated capital is converted to the consumption good in formulating the goods market clearing condition. This is without loss of generality because after production, firms can convert between capital and the consumption good at a one-to-one ratio. Consistent with this convention, in formulating the investment decision, we assume all new capital is converted from the consumption good.

3. *The initial wealth density function $\phi_{i,0}(z)$, following its definition in section 3.5, satisfies*

$$\phi_{i,0}(z) = \int_0^{\infty} a\cdot\Phi_{i,0}(z,a)da.$$

The transition of $\phi_{i,t}(z)$ is implied by the returns and investment decision of the parent firms and the exogenous Markov transition density $f_{i,t}(z'|z)$, and satisfies

$$\phi_{i,t+1}(z') = \int_0^{\infty}\phi_{i,t}(z)\beta\mathbb{E}[R^a_{i,t}(z,\zeta)|z]f_{i,t}(z'|z)dz. \tag{B.3}$$

B.2 Proofs

B.2.1 Lemma 1

Proof. Observing that both the objective and the constraints in the problem are homogeneous of degree one in e_h, we write the problem described in equation (3) as

$$\tilde{R}_{ih}(z,e_h) = \max_{\hat{b}^F_h,\hat{k},\hat{l},\hat{y}}[\hat{y}+(1-\delta)\hat{k}-w_h\hat{l}-(1+r^b_h)\hat{b}^F_h]\cdot e_h, \tag{B.4}$$
$$\begin{aligned} s.t.\quad & \hat{y}\cdot e_h = [\tilde{z}_{ih}(z)\hat{k}]^{\alpha}\hat{l}^{1-\alpha}\cdot e_h \\ & 0\le \hat{b}^F_h\cdot e_h \le \mu_h\cdot e_h \\ & 0\le \hat{k}\cdot e_h\le e_h+\hat{b}^F_h\cdot e_h. \end{aligned}$$

This problem has the same solution as Problem (5), which proves that $\tilde{R}_{ih}(z,e_h) = R_{ih}(z)e_h$.

[6] This condition is equivalent to the following: total production in country h equals the sum of consumption by workers in h, consumption by firm owners in h , investment by firm owners in h, and net outflows.

Define

$$\tilde{\pi}_h(z_{ih}, \hat{k}) \equiv \max_{\tilde{l},\tilde{y}} \tilde{y} - w_h \tilde{l} + (1-\delta)\hat{k}$$
$$s.t. \quad \tilde{y} = [z_{ih}\hat{k}]^{\alpha} \tilde{l}^{1-\alpha}.$$

The first order conditions give the policy functions:

$$\tilde{l}_{ih}(z_{ih}, \hat{k}) = z_{ih}\hat{k}\Big(\frac{1-\alpha}{w_h}\Big)^{1/\alpha},$$
$$\tilde{y}_{ih}(z_{ih}, \hat{k}) = z_{ih}\hat{k}\Big(\frac{1-\alpha}{w_h}\Big)^{(1-\alpha)/\alpha},$$
$$\tilde{\pi}_h(z_{ih}, \hat{k}) = \pi_h(z_{ih})\hat{k}, \text{ where } \pi_h(z_{ih}) = \alpha z_{ih}\Big(\frac{1-\alpha}{w_h}\Big)^{(1-\alpha)/\alpha} + 1 - \delta.$$

Next consider

$$R_{ih}(z) = \max_{\hat{k},\hat{b}^F} \tilde{\pi}_h\big(\tilde{z}_{ih}(z), \hat{k}\big) - (1+r^b_h)\hat{b}^F$$
$$s.t. \quad 0 \le \hat{b}^F_h \le \mu_h$$
$$0 \le \hat{k} \le 1 + \hat{b}^F.$$

Apply $\tilde{\pi}_h(z_{ih}, \hat{k}) = \pi_h(z_{ih})\hat{k}$ that is derived above, we have the policy functions:

$$\hat{b}^F_{ih}(z) = \begin{cases} \mu_h, & \forall \tilde{z}_{ih}(z) \ge z^*_{ih} \\ 0, & \forall \tilde{z}_{ih}(z) < z^*_{ih} \end{cases}$$
$$\hat{k}_{ih}(z) = [1 + \hat{b}^F_{ih}(z)],$$

where z^*_{ih} is determined by that $\pi_h(z^*_{ih}) = (1+r^b_h)$. Under these policy functions

$$R_{ih}(z) = \pi_h\big(\tilde{z}_{ih}(z)\big)[1 + \hat{b}^F_{ih}(z)] - (1+r^b_h)\hat{b}^F_{ih}(z).$$

□

B.2.2 Lemma 2

Proof. We add back the time subscript to be clear that the proof holds for the transition path. We use a guess-and-verify strategy. Suppose the value function in equation (4) takes the following form:

$$v_{i,t}(z, \boldsymbol{\zeta}, a) = \hat{v}_{i,t}(z, \boldsymbol{\zeta}) + B\log(a), \tag{B.5}$$

where $\hat{v}_{i,t}(z,\boldsymbol{\zeta})$ and B are functions and coefficients to be determined. Plug the guess into the right hand side of equation (4), we obtain

$$\begin{aligned} v_{i,t}(z,\boldsymbol{\zeta},a) = & \max_{c,a',\{e_h\}_{h=1}^N,b^H} \log(c) + \beta\mathbb{E}\left[(\hat{v}_{i,t+1}(z',\boldsymbol{\zeta}') + B\log(a'))|z\right] \\ s.t. \quad & \sum_h e_h = a + b^H \\ & -a \le b^H \le \lambda_{i,t}\cdot a \\ & c + a' = \sum_h R_{ih,t}(z)\bar{\eta}_{ih}\zeta_h e_h - (1+r^b_{i,t})b^H. \end{aligned} \tag{B.6}$$

The problem can be solved in two steps. In the first step, firms solve the investment allocation problem by choosing $\{e_h\}_{h=1}^N$ and b^H to maximize total net return on net worth a

$$\begin{aligned} \tilde{R}^a_{i,t}(z,\boldsymbol{\zeta},a) = & \max_{\{e_h\}_{h=1}^N,b^H} \sum_h R_{ih,t}(z)\bar{\eta}_{ih}\zeta_h e_h - (1+r^b_{i,t})b^H \\ s.t. \quad & \sum_h e_h = a + b^H \\ & -a \le b^H \le \lambda_{i,t}\cdot a. \end{aligned}$$

Since the objective is linear in b^H and e_h, if $1+r^b_{i,t} > \max_{h'} R_{ih',t}(z)\bar{\eta}_{ih'}\zeta_{h'}$, the firm stays idle and loans out all the net worth, i.e., $b^H = -a$. If $1+r^b_{i,t} < \max_{h'} R_{ih',t}(z)\bar{\eta}_{ih'}\zeta_{h'}$, the firm stays active, borrows to the maximum, and allocates e_h to hosts that attains $\max_{h'} R_{ih',t}(z)\bar{\eta}_{ih'}\zeta_{h'}$. If $1+r^b_{i,t} = \max_{h'} R_{ih',t}(z)\bar{\eta}_{ih'}\zeta_{h'}$, the firm is indifferent between being idle and active. Therefore, we have $\tilde{R}^a_{i,t}(z,\boldsymbol{\zeta},a) = R^a_{i,t}(z,\boldsymbol{\zeta})\cdot a$ with

$$R^a_{i,t}(z,\boldsymbol{\zeta}) = \begin{cases} \left[\max_{h'} R_{ih',t}(z)\bar{\eta}_{ih'}\zeta_{h'}\right](1+\lambda_{i,t}) - (1+r^b_{i,t})\lambda_{i,t} & \text{if } \max_{h'} R_{ih',t}(z)\bar{\eta}_{ih'}\zeta_{h'} \ge 1+r^b_{i,t} \\ (1+r^b_{i,t}) & \text{if } \max_{h'} R_{ih',t}(z)\bar{\eta}_{ih'}\zeta_{h'} < 1+r^b_{i,t}. \end{cases}$$

The right hand side of equation (B.6) then reduces to

$$\max_{a'} \log(R^a_{i,t}(z,\boldsymbol{\zeta})a - a') + \beta\mathbb{E}\left[(\hat{v}_{i,t+1}(z',\boldsymbol{\zeta}') + B\log(a'))|z\right].$$

Taking first order condition with respect to a', we have

$$a' = \frac{\beta B}{1+\beta B} R^a_{i,t}(z,\boldsymbol{\zeta})a. \tag{B.7}$$

Plug (B.7) into (B.6) and we have

$$v_{i,t}(z,\boldsymbol{\zeta},a) = \log\left(\frac{R^a_{i,t}(z,\boldsymbol{\zeta})}{1+\beta B}\right) + \log(a) + \beta\mathbb{E}[\hat{v}_{i,t+1}(z',\boldsymbol{\zeta}')|z] + \beta B\log\left(\frac{\beta B}{1+\beta B}R^a_{i,t}(z,\boldsymbol{\zeta})a\right). \tag{B.8}$$

Comparing with equation (B.5) for the coefficients to be determined, we can verify that B and $\hat{v}_{i,t}(z,\boldsymbol{\zeta})$

that satisfy the following conditions solve the Bellman equation:

$$
\begin{aligned}
&1+\beta B = B \\
&\hat{v}_{i,t}(z,\boldsymbol{\zeta}) = \log\left(\frac{R^a_{i,t}(z,\boldsymbol{\zeta})}{1+\beta B}\right) + \beta \mathbb{E}[\hat{v}_{i,t+1}(z',\boldsymbol{\zeta}')|z] + \beta B \log\left(\frac{\beta B}{1+\beta B} R^a_{i,t}(z,\boldsymbol{\zeta})\right)
\end{aligned}
\tag{B.9}
$$

This implies the following, which completes the proof of Lemma 2.

$$
\begin{aligned}
B &= \frac{1}{1-\beta}, \\
a' &= \beta R^a_{i,t}(z,\boldsymbol{\zeta}) a \\
c &= R^a_{i,t}(z,\boldsymbol{\zeta}) a - a' = (1-\beta) R^a_{i,t}(z,\boldsymbol{\zeta}) a.
\end{aligned}
$$

□

Solving the value function. In the above, solving the value functions is not necessary because in our benchmark model with log period utility neither $a'(z,\boldsymbol{\zeta},a)$ nor $c(z,\boldsymbol{\zeta},a)$, characterized above, depends on firm values. As values are needed in characterizing firms' decisions in various extensions of the baseline model presented in Section B.6 of this appendix, we discuss how they can be solved.

First, note that solving directly for the fixed point of equation (B.9) requires integrating over the space of $\boldsymbol{\zeta}$. This is in general difficult because the dimension of $\boldsymbol{\zeta}$ is too large. It is feasible under the distribution assumption on $\boldsymbol{\zeta}$ we maintain. Denote $\tilde{v}_{i,t}(z) = \mathbb{E}[\hat{v}_{i,t}(z,\boldsymbol{\zeta})|z]$. Taking conditional expectation on $\boldsymbol{\zeta}$ for equation (B.9) (plugging in $B = \frac{1}{1-\beta}$) gives a functional equation for $\tilde{v}_{i,t}(z)$:

$$
\tilde{v}_{i,t}(z) = \log(1-\beta) + \frac{\beta}{1-\beta}\log(\beta) + \frac{1}{1-\beta}\mathbb{E}[\log(R^a_{i,t}(z,\boldsymbol{\zeta}))|z] + \beta \mathbb{E}[\tilde{v}_{i,t+1}(z')|z], \tag{B.10}
$$

where $\mathbb{E}[\log(R^a_{i,t}(z,\boldsymbol{\zeta}))|z]$ can be characterized in closed form, following a similar strategy in Lemma 3 for $\mathbb{E}[R^a_{i,t}(z,\boldsymbol{\zeta})|z]$. Note that the value function takes only one state variable, z, so it can be solved with standard numerical methods, e.g., via value function iterations, for functionals $\tilde{v}_{i,t}(z)$ and $\mathbb{E}[\tilde{v}_{i,t+1}(z')|z]$. Noting also by the law of iterated expectations, $\mathbb{E}[\hat{v}_{i,t+1}(z',\boldsymbol{\zeta}')|z] = \mathbb{E}[\tilde{v}_{i,t+1}(z')|z]$. We can plug these solutions to equation (B.9) for $\hat{v}_{i,t}(z,\boldsymbol{\zeta})$, and then plug $\hat{v}_{i,t}(z,\boldsymbol{\zeta})$ to equation (B.5) for $v_{i,t}(z,\boldsymbol{\zeta},a)$.

In more general cases, when we deviate from the log utility assumption, or when we allow for switching cost for firms moving production from one country to another, policy functions will depend on $\hat{v}_{i,t}(z,\boldsymbol{\zeta},a)$. In these cases, solving for the value function is necessary. Section B.6 of this appendix shows that the same strategy works for extensions with CRRA utility function or with firm-level switching cost, so much of the tractability in the baseline model carries over.

B.2.3 Lemma 3

Before proving Lemma 3, we first characterize several properties of the correlated Pareto distribution in Lemma B.1. Some of these properties are covered by Arkolakis et al. (2017).

Lemma B.1. *Suppose $\boldsymbol{\zeta} = (\zeta)_{h=1}^N$ follows the standardized correlated Pareto distribution with cumulative distri-*

bution function (CDF):

$$Pr(\zeta_1 \le \tilde{\zeta}_1, ..., \zeta_h \le \tilde{\zeta}_h) = \begin{cases} 1 - \Big(\sum_h \frac{1}{N}\big[\tilde{\zeta}_h^{-\theta}\big]^{\frac{1}{1-\rho}}\Big)^{1-\rho}, & \textit{if } \forall h,\ \tilde{\zeta}_h \ge 1, \\ 0, & \textit{if } \exists h \textit{ s.t. } \tilde{\zeta}_h < 1. \end{cases}$$

Define $\Xi \equiv \max_{h'} A_{h'}\zeta_{h'}$, *in which* $A_h,\ h = 1, ..., N$ *are any positive constants.* Ξ *has the following properties:*

1. *The CDF for* Ξ *is*

$$Pr(\Xi \le B) = \begin{cases} 1 - \tilde{A}^{\theta} B^{-\theta}, & \textit{if } B \ge \max_{h'} A_{h'} \\ 0, & \textit{if } B < \max_{h'} A_{h'}. \end{cases}$$

   where $\tilde{A} = \Big(\frac{1}{N}\sum_{h'} A_{h'}^{\frac{\theta}{1-\rho}}\Big)^{\frac{1-\rho}{\theta}}$. *This immediately implies*

$$Pr(\Xi = \max_{h'} A_{h'}) = 1 - \tilde{A}^{\theta}[\max_{h'} A_{h'}]^{-\theta}.$$

2. *The bottom-truncated conditional mean of* Ξ *is*

$$\mathbb{E}[\Xi|\Xi \ge B] = \begin{cases} \frac{\theta}{\theta-1}B, & \textit{if } B > \max_{h'} A_{h'} \\ (1 - \tilde{A}^{\theta}[\max_{h'} A_{h'}]^{-\theta}) \max_{h'} A_{h'} + \tilde{A}^{\theta}[\max_{h'} A_{h'}]^{-\theta} \frac{\theta}{\theta-1} \max_{h'} A_{h'}, & \textit{if } B \le \max_{h'} A_{h'} \end{cases}$$

3. *The conditional probability of* Ξ *being achieved at host h, under two different conditions, is:*

$$\forall B > \max_{h'} A_{h'}\zeta_{h'},\ Pr(\arg\max_{h'} A_{h'}\zeta_{h'} = h \Big| \Xi \ge B) = \frac{A_h^{\theta/(1-\rho)}}{\sum_{h'} A_{h'}^{\theta/(1-\rho)}}$$

   If the set $\overline{\mathbb{H}} \equiv \arg\max_{h'} A_{h'}$ *is a singleton, then*

$$Pr(\arg\max_{h'} A_{h'}\zeta_{h'} = \bar{h} \Big| \Xi = \max_{h'} A_{h'}) = \begin{cases} 1, & \textit{if } \bar{h} \in \overline{\mathbb{H}} \\ 0, & \textit{if } \bar{h} \notin \overline{\mathbb{H}} \end{cases}$$

Proof. 1. Consider

$$\begin{aligned} Pr(\Xi \le B) &= Pr(A_1\zeta_1 \le B, ..., A_h\zeta_h \le B) \\ &= Pr(\zeta_1 \le \frac{B}{A_1}, ..., \zeta_h \le \frac{B}{A_h}) \\ &= 1 - \Big(\sum_{h'} \frac{1}{N}\big[(\frac{B}{A_{h'}})^{-\theta}\big]^{\frac{1}{1-\rho}}\Big)^{1-\rho}, \quad \text{for } \frac{B}{A_{h'}} \ge 1, \forall h' \\ &= 1 - \Big(\frac{1}{N}\sum_{h'} A_{h'}^{\frac{\theta}{1-\rho}}\Big)^{1-\rho} B^{-\theta}, \quad \text{for } B \ge \max_{h'} A_{h'} \end{aligned}$$

If $B < \max_{h'} A_{h'}$, then $\exists h$ s.t. $\frac{B}{A_h} < 1$. Therefore,

$$Pr(\Xi \leq B) = Pr(\zeta_1 \leq \frac{B}{A_1}, ..., \zeta_h \leq \frac{B}{A_h}) = 0.$$

Therefore,

$$Pr(\Xi = \max_{h'} A_{h'}) = Pr(\Xi \leq \max_{h'} A_{h'}) - \lim_{B \uparrow \max_{h'} A_{h'}} Pr(\Xi \leq B) = 1 - \Big(\frac{1}{N}\sum_{h'} A_{h'}^{\frac{\theta}{1-\rho}}\Big)^{1-\rho} [\max_{h'} A_{h'}]^{-\theta}.$$

2. For $B > \max_{h'} A_{h'}$, from part 1, $\forall C \geq B$

$$Pr(\Xi > C | \Xi > B) = \Big(\frac{C}{B}\Big)^{-\theta}.$$

Therefore, $\Xi | \Xi > B$ follows a Pareto distribution with tail parameter θ and scale parameter B. Thus we have

$$\mathbb{E}[\Xi | \Xi > B] = \frac{\theta}{\theta - 1} B.$$

Because $Pr(\Xi \leq B)$ is continuous in B when $B > \max_{h'} A_{h'}$, we have $Pr(\Xi = B) = 0$. Therefore:

$$\mathbb{E}[\Xi | \Xi \geq B] = \mathbb{E}[\Xi | \Xi > B] = \frac{\theta}{\theta - 1} B.$$

For $B \leq \max_{h'} A_{h'}$, since $Pr(\Xi \geq B) = 1$, we have

$$\begin{aligned}
\mathbb{E}[\Xi | \Xi \geq B] &= \mathbb{E}(\Xi) \\
&= Pr(\Xi = \max_{h'} A_{h'}) \mathbb{E}(\Xi | \Xi = \max_{h'} A_{h'}) + Pr(\Xi > \max_{h'} A_{h'}) \mathbb{E}(\Xi | \Xi > \max_{h'} A_{h'}) \\
&= (1 - \tilde{A}^{\theta} [\max_{h'} A_{h'}]^{-\theta}) \max_{h'} A_{h'} + \tilde{A}^{\theta} [\max_{h'} A_{h'}]^{-\theta} \frac{\theta}{\theta - 1} \max_{h'} A_{h'}.
\end{aligned}$$

3. For $B > \max_{h'} A_{h'}$,

$$Pr(\arg\max_{h'} A_{h'}\zeta_{h'} = h \wedge \Xi \geq B) = \int_B^{\infty} Pr(A_{h'}\zeta_{h'} \leq u, \forall h' \neq h \big| A_h \zeta_h = u) g_h(u) du,$$

where $g_h(u)$ is the marginal distribution of $A_h \zeta_h$.

For $u \geq B > \max_{h'} A_{h'}$, the integrand is an explicit function of u:

$$\begin{aligned}
Pr(A_{h'}\zeta_{h'} \leq u, \forall h' \neq h \big| A_h \zeta_h = u) g_h(u) &= \frac{\partial Pr(A_1\zeta_1 \leq u, A_h\zeta_h \leq C, A_{h'}\zeta_{h'} \leq u)}{\partial C}\Big|_{C=u} \\
&= \frac{A_h^{\frac{\theta}{1-\rho}}}{N} \Big(\frac{1}{N}\sum_{h'} A_{h'}^{\frac{\theta}{1-\rho}}\Big)^{-\rho} \theta u^{-\theta-1}.
\end{aligned}$$

Therefore,

$$Pr(\arg\max_{h'} A_{h'}\zeta_{h'} = h \wedge \Xi \geq B) = \frac{A_h^{\frac{\theta}{1-\rho}}}{N}\left(\frac{1}{N}\sum_{h'} A_{h'}^{\frac{\theta}{1-\rho}}\right)^{-\rho} B^{-\theta}.$$

And

$$\begin{aligned}
Pr(\arg\max_{h'} A_{h'}\zeta_{h'} = h \Big| \Xi \geq B) &= \frac{Pr(\arg\max_{h'} A_{h'}\zeta_{h'} = h \wedge \Xi \geq B)}{Pr(\Xi \geq B)} \\
&= \frac{\frac{A_h^{\frac{\theta}{1-\rho}}}{N}\left(\frac{1}{N}\sum_{h'} A_{h'}^{\frac{\theta}{1-\rho}}\right)^{-\rho} B^{-\theta}}{\left(\frac{1}{N}\sum_{h'} A_{h'}^{\frac{\theta}{1-\rho}}\right)^{1-\rho} B^{-\theta}} \\
&= \frac{A_h^{\frac{\theta}{1-\rho}}}{\sum_{h'} A_{h'}^{\frac{\theta}{1-\rho}}},
\end{aligned}$$

where the second equality applies $Pr(\Xi \geq B) = \left(\frac{1}{N}\sum_{h'} A_{h'}^{\frac{\theta}{1-\rho}}\right)^{1-\rho} B^{-\theta}$ from part 1 of the lemma.

If $\overline{\mathbb{H}} = \arg\max_{h'} A_{h'}$ is a singleton, then $\forall h \notin \overline{\mathbb{H}}$, we have $A_h < \max_{h'} A_{h'}$, and

$$Pr\Big(A_h\zeta_h = \max_{h'} A_{h'} \ \wedge \ \Xi = \max_{h'} A_{h'}\Big) \leq Pr\Big(\zeta_h = \frac{\max_{h'} A_{h'}}{A_h}\Big) = 0, \tag{B.11}$$

in which the equality follows from $\frac{\max_{h'} A_{h'}}{A_h} > 1$ and that

$$\lim_{x\to\infty} Pr(\zeta_1 \leq x, ..., \zeta_h \leq \tilde{\zeta}_h, ...\zeta_N \leq x)$$

is a continuous function of $\tilde{\zeta}_h$ for $\tilde{\zeta}_h > 1$.

Note also

$$\begin{aligned}
Pr\Big(\Xi = \max_{h'} A_{h'}\Big) &\leq \sum_{\tilde{h}} Pr\Big(A_{\tilde{h}}\zeta_{\tilde{h}} = \max_{h'} A_{h'} \wedge \ \Xi = \max_{h'} A_{h'}\Big) \\
&= \sum_{\tilde{h}\in\overline{\mathbb{H}}} Pr\Big(A_{\tilde{h}}\zeta_{\tilde{h}} = \max_{h'} A_{h'} \wedge \ \Xi = \max_{h'} A_{h'}\Big) \\
&= Pr\Big(A_{\bar{h}}\zeta_{\bar{h}} = \max_{h'} A_{h'} \wedge \ \Xi = \max_{h'} A_{h'}\Big), \text{ for } \bar{h} \in \overline{\mathbb{H}},
\end{aligned}$$

in which the first line follows the rule of total probability, the second line follows from equation (B.11), and the third line follows from the fact that $\overline{\mathbb{H}}$ is a singleton.

By definition, the following is also true

$$Pr\Big(A_{\bar{h}}\zeta_{\bar{h}} = \max_{h'} A_{h'} \wedge \ \Xi = \max_{h'} A_{h'}\Big) \leq Pr\Big(\Xi = \max_{h'} A_{h'}\Big),$$

We thus have

$$Pr\Big(A_{\bar{h}}\zeta_{\bar{h}} = \max_{h'} A_{h'} \wedge \Xi = \max_{h'} A_{h'}\Big) = Pr\Big(\Xi = \max_{h'} A_{h'}\Big),$$

which also implies

$$Pr\Big(\arg\max_{h'} A_{h'}\zeta_{h'} = \bar{h}\Big|\Xi = \max_{h'} A_{h'}\Big) = 1.$$

$\square$

Proof of Lemma 3.

Proof. We omit the time subscript. The proof below uses a special case of Lemma B.1 with $\rho = 0$.

1. Define $\Xi = \max_{h'} \bar{\eta}_{ih'} R_{ih'}(z)\zeta_{h'}$. Notice we have defined $\overline{R}_i(z) \equiv \max_{h'} \bar{\eta}_{ih'} R_{ih'}(z)$ and $\tilde{R}_i(z) \equiv \Big(\frac{1}{N}\sum_{h'} [\bar{\eta}_{ih'} R_{ih'}(z)]^\theta\Big)^{\frac{1}{\theta}}$. If $1 + r_i^b > \overline{R}_i(z)$, applying part 1 of Lemma B.1, we have

$$Pr(\Xi \geq 1 + r_i^b | z) = [\tilde{R}_i(z)]^\theta (1 + r_i^b)^{-\theta}.$$

From part 3 of Lemma B.1, we have the probability of investing in h conditional on being active

$$\chi_{ih}(z) = \frac{1}{N}\Big(\frac{\bar{\eta}_{ih} R_{ih}(z)}{\tilde{R}_i(z)}\Big)^\theta, \tag{B.12}$$

and the unconditional probability

$$e_{ih}(z) = [\tilde{R}_i(z)]^\theta (1 + r_i^b)^{-\theta} \frac{1}{N}\Big(\frac{\bar{\eta}_{ih} R_{ih}(z)}{\tilde{R}_i(z)}\Big)^\theta,$$

Applying the definition in equation (7) and part 2 of Lemma B.1, we can derive:

$$\mathbb{E}[R_i^a(z, \boldsymbol{\zeta})|z] = \underbrace{\Big(1 - [\tilde{R}_i(z)/(1 + r_i^b)]^\theta\Big)}_{Pr(\Xi < 1 + r_i^b|z)}(1 + r_i^b) + \underbrace{[\tilde{R}_i(z)/(1 + r_i^b)]^\theta}_{Pr(\Xi \geq 1 + r_i^b|z)}\Big(\underbrace{\frac{\theta}{\theta - 1}(1 + r_i^b)}_{\mathbb{E}[\Xi|\Xi \geq 1 + r_i^b|z]}\overbrace{(1 + \lambda_i) - (1 + r_i^b)\lambda_i}^{\text{leverage adjustment}}\Big).$$

This proves part **(i)**.

2. For part **(ii)**, if $1 + r_i^b < \overline{R}_i(z)$, from part 1 of Lemma B.1

$$Pr(\Xi \geq 1 + r_i^b | z) = 1,$$

i.e., all firms with productivity z are active.

If set $\overline{\mathbb{H}} \equiv \arg\max_{h'} \bar{\eta}_{ih'} R_{ih'}(z)$ is a singleton, applying part 3 of Lemma B.1 gives

$$Pr\Big(\arg\max_{h'} \bar{\eta}_{ih'} R_{ih'}(z)\zeta_{h'} = \bar{h}\Big|\Xi = \overline{R}_i(z), z\Big) = 1,$$

which says that conditional on $\max_{h'} \bar{\eta}_{ih'} R_{ih'}(z)\zeta_{h'} = \overline{R}_i(z)$, with probability one the investment goes to the host $\bar{h} \in \overline{\mathbb{H}}$. The share of firms (all of which active) investing in country h is:

For $h \in \overline{\mathbb{H}}$

$$\begin{aligned}\hat{e}_{ih}(z) &= \underbrace{\left(1 - [\tilde{R}_i(z)/\overline{R}_i(z)]^\theta\right)}_{Pr(\Xi = \overline{R}_i(z)|z)} + \underbrace{[\tilde{R}_i(z)/\overline{R}_i(z)]^\theta}_{Pr(\Xi > \overline{R}_i(z)|z)} \cdot \chi_{ih}(z) \\ &= 1 - [1 - \chi_{ih}(z)][\tilde{R}_i(z)/\overline{R}_i(z)]^\theta, \text{ with } \chi_{ih}(z) \text{ defined in equation (B.12);}\end{aligned}$$

For $h \notin \overline{\mathbb{H}}$,

$$\hat{e}_{ih}(z) = \chi_{ih}(z)[\tilde{R}_i(z)/\overline{R}_i(z)]^\theta.$$

From part 2 of Lemma B.1, the return conditional on z is

$$\mathbb{E}[R_i^a(z, \boldsymbol{\zeta})|z] = \underbrace{\left(1 - [\tilde{R}_i(z)/\overline{R}_i(z)]^\theta\right)}_{Pr(\Xi = \overline{R}_i(z)|z)} \overline{R}_i(z)(1+\lambda_i) + \underbrace{[\tilde{R}_i(z)/\overline{R}_i(z)]^\theta}_{Pr(\Xi > \overline{R}_i(z)|z)} \left(\underbrace{\frac{\theta}{\theta - 1}\overline{R}_i(z)}_{\mathbb{E}[\Xi|\Xi > \overline{R}_i(z)|z]} (1+\lambda_i)\right) - (1 + r_i^b)\lambda_i.$$

$\square$

In establishing the second part of the lemma, we explicitly consider the firms whose realizations of $\Xi(z)$ are exactly at the mass point $\overline{R}_i(z)$. This part of the lemma maintains the assumption that $\overline{R}_i(z) \equiv \max_{h'} \bar{\eta}_{ih'} R_{ih'}(z)$ is achieved by only one host.[7]

B.2.4 Equation (9)

Lemma B.2. *The wealth density functions satisfy*

$$\phi_{i,t+1}(z') = \int_0^\infty \phi_{i,t}(z)\beta\mathbb{E}[R_{i,t}^a(z, \boldsymbol{\zeta})|z] f_{i,t}(z'|z) dz.$$

Proof. Without loss of generality, assume the mass of firms in country i to be 1 and index firm by $\omega \in [0, 1]$. Accordingly, denote the net worth of firm ω in country i at time t by $a_{i,t}(\omega)$ and the productivity

[7] When the maximum return is achieved in more than one country, firms at the mass point will be indifferent between hosts, in which case a tie-breaking rule is in principle needed. In the quantification, however, such a choice does not matter. Unless two hosts h and h' have the exact same primitives and equilibrium wages, $\bar{\eta}_{ih'} R_{ih'}(z)$ can cross $\bar{\eta}_{ih} R_{ih}(z), h \neq h'$ for only finitely many values of z. Because we specify z to have a continuous density, outside the special case with two identical host countries, $\max_{h'} \bar{\eta}_{ih'} R_{ih'}(z)$ can be achieved by more than one country on only a zero-measure set.

by $z_{i,t}(\omega)$. Then consider the CDF of $\phi_{i,t+1}(\cdot)$, denoted by $\Psi_{i,t+1}(\cdot)$:

$$
\begin{aligned}
\Psi_{i,t+1}(z') &\equiv \int_0^{z'} \phi_{i,t+1}(\tilde{z})\, d\tilde{z} \\
&= \int_0^1 a_{i,t+1}(\omega)\mathbb{1}\left(z_{i,t+1}(\omega) \leq z'\right) d\omega \\
&= \int_0^1 \int_0^{z'} a_{i,t+1}(\omega) f_{i,t}(\tilde{z}|z_{i,t}(\omega)) d\tilde{z}\, d\omega \\
&= \int_0^{z'} \int_0^1 a_{i,t}(\omega) \cdot \beta \mathbb{E}[R^a_{i,t}(z_{i,t}(\omega), \boldsymbol{\zeta})|z_{i,t}(\omega)] \cdot f_{i,t}(\tilde{z}|z_{i,t}(\omega)) d\omega\, d\tilde{z} \\
&= \int_0^{z'} \left[\int_0^\infty \int_0^\infty a \cdot \beta \mathbb{E}[R^a_{i,t}(z, \boldsymbol{\zeta})|z] f_{i,t}(\tilde{z}|z) \Phi_{i,t}(z,a) da\, dz \right] d\tilde{z} \\
&= \int_0^{z'} \left[\int_0^\infty \beta \mathbb{E}[R^a_{i,t}(z, \boldsymbol{\zeta})|z] f_{i,t}(\tilde{z}|z) \int_0^\infty a \cdot \Phi_{i,t}(z,a) da\, dz \right] d\tilde{z} \\
&= \int_0^{z'} \left[\int_0^\infty \phi_{i,t}(z) \beta \mathbb{E}[R^a_{i,t}(z, \boldsymbol{\zeta})|z] f_{i,t}(\tilde{z}|z) dz \right] d\tilde{z},
\end{aligned}
$$

in which we have applied the definition of $\phi_{i,t+1}$, the definition of $\phi_{i,t}(z)$, and the Fubini-Tonelli Theorem. Differentiating $\Psi_{i,t+1}(z')$ with respect to z' gives $\phi_{i,t+1}(z') = \Psi'_{i,t+1}(z') = \int_0^\infty \phi_{i,t}(z)\beta\mathbb{E}[R^a_{i,t}(z,\boldsymbol{\zeta})|z] f_{i,t}(z'|z)dz$. □

B.2.5 Proposition 1

Proof. We consider the empirically relevant case of $\overline{R}_i(z) \equiv \max_{h'} \bar{\eta}_{ih'} R_{ih'}(z) = R_{ii}(z)$, i.e., firms prefer investing domestically to investing abroad when all idisocynratic draws from all hosts are the same. This also means for all firms, investing domestically offers the highest expected return. We verify that this case holds when the return wedges are calibrated to match bilateral FDI.

Under this scenario, the policy function characterized in Lemma 3 can be written as:

$$
\hat{e}_{ih}(z) = \frac{1}{N} \bar{\eta}^\theta_{ih} \left(\frac{R_{ih}(z)}{\overline{\overline{R}}_i(z)} \right)^\theta, \tag{B.13}
$$

where $\overline{\overline{R}}_i(z) \equiv \max\{1 + r^b_i, R_{ii}(z)\}$ determines which scenario of Lemma (3) applies. We can then write

$$
\begin{aligned}
[FDI]_{ih} &= \int_0^\infty \psi_{ih}(z) dz = W_i \int_0^\infty (1+\lambda_i) \hat{e}_{ih}(z) \hat{\phi}_i(z) dz \\
&= W_i (1+\lambda_i) \int_0^\infty \frac{1}{N} \bar{\eta}^\theta_{ih} \left(\frac{R_{ih}(z)}{\overline{\overline{R}}_i(z)} \right)^\theta \hat{\phi}_i(z) dz \\
&= \frac{1}{N} W_i (1+\lambda_i) \bar{\eta}^\theta_{ih} \left(\frac{R_{ih}(\overline{\overline{z}}_i)}{R_{ii}(\overline{\overline{z}}_i)} \right)^\theta [\epsilon^{FDI}_{ih}]^\theta \\
\text{where } \epsilon^{FDI}_{ih} &\equiv \left[\int_0^\infty \underbrace{\hat{\phi}_i(z)}_{\text{equity weighted}} \Bigg(\underbrace{\frac{R_{ih}(z)}{\overline{\overline{R}}_i(z)} \Big/ \frac{R_{ih}(\overline{\overline{z}}_i)}{R_{ii}(\overline{\overline{z}}_i)}}_{\text{geometric difference in (1+ROE) relative to the average firm}} \Bigg)^\theta dz \right]^{\frac{1}{\theta}}.
\end{aligned}
$$

Note that $\epsilon_{ih}^{FDI} = 1$ if there is no firm heterogeneity (i.e., $z = \overline{\overline{z}}_i$).[8] Next, apply the policy function of $\hat{y}_{ih}(z)$ that is characterized in Lemma 1, consider

$$\begin{aligned}
Y_{ih} &= \int_0^\infty \hat{y}_{ih}(z)\psi_{ih}(z)dz \\
&= [FDI]_{ih} \int_0^\infty \tilde{z}_{ih}(z)[1+\hat{b}_{ih}^F(z)]\left(\frac{1-\alpha}{w_h}\right)^{(1-\alpha)/\alpha} \frac{\psi_{ih}(z)}{\int_0^\infty \psi_{ih}(z')dz'}dz \\
&= [FDI]_{ih} \cdot \tilde{z}_{ih}(\overline{\overline{z}}_i) \cdot \left(\frac{1-\alpha}{w_h}\right)^{(1-\alpha)/\alpha} \underbrace{\frac{\int_0^\infty \psi_{ih}(z')[1+\hat{b}_{ih}^F(z')]dz'}{\int_0^\infty \psi_{ih}(z')dz'}}_{1+\overline{lev}_{ih}^F} \underbrace{\int_0^\infty \frac{\psi_{ih}(z)[1+\hat{b}_{ih}^F(z)]}{\int_0^\infty \psi_{ih}(z')[1+\hat{b}_{ih}^F(z')]dz'} \frac{\tilde{z}_{ih}(z)}{\tilde{z}_{ih}(\overline{\overline{z}}_i)}dz}_{\epsilon_{ih}^Y,\text{ asset weighted average geometric difference in productivity}}.
\end{aligned}$$

□

B.2.6 Proposition 2

Proof. Following the definition in Section 3.5, the sum of equity investment in host country h by parents with productivity z from home country i is:

$$\psi_{ih}(z) = (1+\lambda_i)\hat{e}_{ih}(z)\phi_i(z).$$

Define $y(w_h) = \left(\frac{1-\alpha}{w_h}\right)^{(1-\alpha)/\alpha}$ as the output from each unit of capital used by an affiliate whose productivity is one and $l(w_h) = \left(\frac{1-\alpha}{w_h}\right)^{1/\alpha}$ as the labor used by that affiliate in producing $y(w_h)$. Because firms from country h are restricted from investing overseas, their investment decision reduces to a threshold rule: a firm leverages up and invests domestically if $z > z_h^*$, and stays idle if $z < z_h^*$, where z_h^* is determined by

$$\alpha z_h^* y(w_h) = r_h^b + \delta. \tag{B.14}$$

Using the policy function, the aggregate wealth and the wealth share density function defined in Section 3.5, the total capital used by domestic firms is

$$W_h \int_{z_h^*}^\infty \hat{\phi}_h(z)(1+\lambda_h)dz = K_{hh}. \tag{B.15}$$

The labor market clearing condition is given by:

$$\left[\sum_{i\neq h}\int_0^\infty \psi_{ih}(z)[1+\hat{b}_{ih}^F(z)]\tilde{z}_{ih}(z)dz + W_h\int_{z_h^*}^\infty \hat{\phi}_h(z)(1+\lambda_h)\tilde{z}_{hh}(z)dz\right] l(w_h) = L_h,$$

Applying the normalization $\tilde{z}_{hh}(z) = z$ and noticing that the production in country h by affiliates from

[8] Note that $\overline{\overline{R}}_i(\overline{\overline{z}}_i) = R_{ii}(\overline{\overline{z}}_i)$ as long as the firm with the average productivity $\overline{\overline{z}}_i$ is active.

country i is $Y_{ih} = y(w_h) \int_0^\infty \psi_{ih}(z)[1 + \hat{b}_{ih}^F(z)]\tilde{z}_{ih}(z)dz$, we rewrite the labor market clearing condition as:

$$\left[\sum_{i \neq h} Y_{ih} + y(w_h) W_h \int_{z_h^*}^\infty \hat{\phi}_h(z)(1+\lambda_h) z dz\right] \frac{l(w_h)}{y(w_h)} = L_h. \tag{B.16}$$

Applying that z follows the Pareto distribution, $\hat{\phi}_h(z) = \gamma \bar{z}_h^\gamma z^{-1-\gamma}$ for $z \geq \bar{z}_h$, where $\bar{z}_h$ is the location parameter f the Pareto distribution, and focusing on the interesting case that $z_h^* > \bar{z}_h$, we have

$$\begin{aligned} \int_{z_h^*}^\infty \hat{\phi}_h(z) \cdot z \cdot dz &= \frac{\gamma}{\gamma - 1} \bar{z}_h^\gamma [z_h^*]^{1-\gamma} \\ \int_{z_h^*}^\infty \hat{\phi}_h(z) dz &= \bar{z}_h^\gamma [z_h^*]^{-\gamma} \end{aligned} \tag{B.17}$$

Combine (B.15), (B.16) and (B.17), we obtain

$$\bar{z}_h^\gamma [z_h^*]^{-\gamma}(1+\lambda_h) = \frac{K_{hh}}{W_h}, \tag{B.18}$$

and

$$\frac{W_h(1+\lambda_h)\frac{\gamma}{\gamma-1}\left(\frac{\frac{K_{hh}}{W_h}}{1+\lambda_h}\right)^{\frac{\gamma-1}{\gamma}} \bar{z}_h}{\frac{Y_{hh}}{Y_h}} l(w_h) = L_h. \tag{B.19}$$

Substituting $l(w_h)$ to equation (B.19) and taking log of both sides, we have

$$Cons + \log(W_h) - \frac{1}{\alpha}\log(w_h) + \frac{\gamma - 1}{\gamma}\log\left(\frac{K_{hh}}{W_h}\right) - \log\left(\frac{Y_{hh}}{Y_h}\right) = 0,$$

where $Cons$ is a constant. Taking the difference between two equilibria with different degree of openness, noting openness does not affect contemporary W_h, gives us:

$$\Delta \log(w_h) = -\alpha \Delta \log\left(\frac{Y_{hh}}{Y_h}\right) + \alpha \frac{\gamma - 1}{\gamma} \Delta \log\left(\frac{K_{hh}}{W_h}\right). \tag{B.20}$$

We can further express the change in capital uses between two equilibria as a function of prices by combining equations (B.14) and (B.18):

$$\Delta \log\left(\frac{K_{hh}}{W_h}\right) = -\gamma \Delta \log(r_h^b + \delta) - \frac{\gamma(1-\alpha)}{\alpha} \Delta \log(w_h),$$

in which $\Delta \log(r_h^b + \delta)$ is the change between the two equilibria in the world interest rate gross of depreciation (or the change in country h's interest rate in the case of segregated bond markets). Plugging this into equation (B.20) gives us equation (13) in the text. □

B.2.7 Proposition 3

We prove the Proposition using four lemmas, stated and proved below. We denote the focal country in Proposition 3 by i and all others by h. We denote the wealth in country i or h at the end of period 1 by $\mathcal{W}_{h,1}$ and $\mathcal{W}_{i,1}$, using the script $\mathcal{W}$ to differentiate them from the beginning-of-period wealth, W. We denote the common interest rate in the economy r^b, dropping the country-specific subscript.

Lemma B.3. *For the two-period economy, the demand of country h's labor by country h's firms at period 2 can be written as*

$$L_{h,2}^H = \mathcal{W}_{h,1}\mathcal{L}^H(\log w_{h,2};\ \Theta_h),$$

where $\Theta_h \equiv (\beta_h, \lambda_{h,2}, \bar{z}_{h,2})$ is the vector of parameters we focus on in the comparative statics; $\mathcal{W}_{h,1}$ is country h's end-of-period-one aggregate net worth, which depends on parameters in period 1 only; $\mathcal{L}^H(\log w_{h,2};\ \Theta_h)$ is independent of $\mathcal{W}_{h,1}$, decreasing and twice differentiable in $\log w_{h,2}$ and increasing in each element of Θ_h.

Proof. Aggregating the labor demand across country h's firms at period 2 gives:

$$L_{h,2}^H = \Big(W_{h,2}\int_{z_{h,2}^*}^{\infty}(1+\lambda_{h,2})z\hat{\phi}_{h,2}(z)dz\Big)\hat{l}(w_{h,2}),$$

where $\hat{l}(w_h) = \left(\frac{1-\alpha}{w_h}\right)^{1/\alpha}$, $z_{h,2}^* = \frac{\bar{r}^b+\delta}{\alpha}\left(\frac{w_{h,2}}{1-\alpha}\right)^{\frac{1-\alpha}{\alpha}}$, $\hat{\phi}_{h,2}(z) = \gamma\bar{z}_{h,2}^{\gamma}z^{-1-\gamma}$.

Apply the policy function of saving we have $W_{h,2} = \beta_h\mathcal{W}_{h,1}$. Therefore,

$$L_{h,2}^H = \mathcal{W}_{h,1}\cdot\underbrace{\frac{\gamma}{\gamma-1}\Big(\frac{\bar{r}^b+\delta}{\alpha}\Big)^{1-\gamma}(1-\alpha)^{[-\frac{(1-\alpha)(1-\gamma)}{\alpha}+\frac{1}{\alpha}]}\cdot w_{h,2}^{\frac{\alpha\gamma-\alpha-\gamma}{\alpha}}\cdot\beta_h(1+\lambda_{h,2})\bar{z}_h^{\gamma}}_{\mathcal{L}^H(\log w_{h,2};\ \Theta_h)},$$

which increases in each element of $\Theta_h \equiv (\beta_h, \lambda_{h,2}, \bar{z}_{h,2})$. $\square$

Lemma B.4. *For the two-period economy, the total demand of country h's labor by firms from country other than h at period 2 can be written as*

$$L_{h,2}^F = \mathcal{W}_{i,1}\mathcal{L}^F(\log w_{h,2}, \Theta_i),$$

where Θ_i is a vector of fundamental parameters of interest, i.e., $\Theta_i \equiv (\beta_i, \lambda_{i,2}, \bar{z}_{i,2}, \bar{\eta}_{ih,2})$. $\mathcal{W}_{i,1}$ is country i's end-of-period-one aggregate net worth, which depends on parameters in period 1 only; $\mathcal{L}^F(\log w_{h,2}, \Theta_i)$ is decreasing and twice differential in $\log w_{h,2}$ and increasing in each element of Θ_i.

Proof. Aggregating the demand of country h's labor across firms of a country $i \neq h$ at period 2, and applying the policy functions that are characterized in Lemma 1-3 gives:

$$L_{ih,2} = \Big(W_{i,2}\int_0^{\infty}(1+\lambda_{i,2})\frac{1}{N}\frac{[\bar{\eta}_{ih,2}\mathcal{R}_{ih,2}(z;w_{h,2})]^{\theta}}{[\overline{\overline{R}}_{i,2}(z)]^{\theta}}z\hat{\phi}_{i,2}(z)dz\Big)\hat{l}(w_{h,2}),$$

where

$$\hat{l}(w_{h,2}) \equiv \Big(\frac{1-\alpha}{w_{h,2}}\Big)^{1/\alpha}$$
$$\mathcal{R}_{ih,2}(z; w_{h,2}) \equiv \Big[\alpha\tilde{z}_{ih}(z)\Big(\frac{1-\alpha}{w_{h,2}}\Big)^{\frac{1-\alpha}{\alpha}} + 1 - \delta\Big]\hat{k}_{ih,2}(z) - (1+r^b)\hat{b}_{ih,2}(z)$$
$$\overline{\overline{R}}_{i,2}(z) \equiv \max\{1+r^b, \overline{R}_{i,2}(z)\}$$
$$\overline{R}_{i,2}(z) = \max_{i' \neq h}\{\bar{\eta}_{ii'} R_{ii',2}(z)\}$$
$$R_{ii',2}(z) = \Big[\alpha\tilde{z}_{ii'}(z)\Big(\frac{1-\alpha}{w_{i',2}}\Big)^{\frac{1-\alpha}{\alpha}} + 1 - \delta\Big]\hat{k}_{ii',2}(z) - (1+r^b)\hat{b}_{ii',2}(z).$$

Note that in these expressions, $\hat{k}_{ih,2}(z), \hat{b}_{ih,2}(z), R_{ii',2}(z)$ are characterized in Lemma 1. As we have assumed $\bar{r}^b$ and $w_{i',2}, i' \neq h$ to be constant, all these expressions are only a function of $\log w_{h,2}$ and fundamental parameters, which means $\log w_{h,2}$ is the only endogenous variable that enters $\mathcal{L}^F$.

Apply the policy function of saving that is characterized in Lemma 2 we have

$$W_{i,2} = \beta_i \mathcal{W}_{i,1},$$

where $\mathcal{W}_{i,1}$ is country i's end-of-period-one aggregate net worth that depends on parameters in period 1 only. Therefore,

$$\begin{aligned} L^F_{h,2} &= \sum_{i \neq h} L_{ih,2} \\ &= \mathcal{W}_{i,1} \cdot \underbrace{\frac{(N-1)}{N}(1-\alpha)^{\frac{1}{\alpha}} \cdot \beta_i(1+\lambda_{i,2}) \cdot w_{h,2}^{-\frac{1}{\alpha}} \cdot \int_0^\infty \frac{[\bar{\eta}_{ih,2}\mathcal{R}_{ih,2}(z; w_{h,2})]^\theta}{[\overline{\overline{R}}_{i,2}(z)]^\theta} z\hat{\phi}_{i,2}(z) dz}_{\mathcal{L}^F(\log w_{h,2};\ \Theta_i)}. \end{aligned}$$

Observe again that the only endogenous variable inside the integration is $w_{h,2}$. Thus, holding $w_{h,2}$ constant, $\mathcal{L}^F(\log w_{h,2};\ \Theta_i)$ increases in each element of $\Theta_i \equiv (\beta_i, \lambda_{i,2}, \bar{z}_{i,2}, \bar{\eta}_{ih,2})$.

□

Lemma B.5. *In the two-period economy, $\mathcal{W}_{h,1}$ and $\mathcal{W}_{i,1}$ depend on parameters at period 1 only, $\frac{d \log \mathcal{W}_{h,1}}{d\bar{\eta}_{ih,1}} < 0$, $\frac{d \log \mathcal{W}_{i,1}}{d\bar{\eta}_{ih,1}} > 0$, and $\frac{d \log \mathcal{W}_{h,1}}{d\bar{\eta}_{ih,1}}$, $\frac{d \log \mathcal{W}_{i,1}}{d\bar{\eta}_{ih,1}}$ are continuous in $\bar{\eta}_{ih,1}$.*

Proof. It follows immediately from Lemma 2, which implies that period 1's operation decisions maximize the static return and do not respond to future parameters. □

Lemma B.6. *In the two-period economy, the equilibrium domestic labor demand $\mathcal{W}_{h,1}\mathcal{L}^H(\log w^*_{h,2};\ \Theta_h)$ defined in Lemma 2 is increasing in each element of Θ_h; the equilibrium foreign labor demand $\mathcal{W}_{i,1}\mathcal{L}^F(\log w^*_{h,2};\ \Theta_i)$ defined in Lemma 3 is increasing in each element of Θ_i. Moreover, the equilibrium labor demands satisfy:*

$$\lim_{x\to\infty}\frac{\mathcal{W}_{h,1}\mathcal{L}^H(\log w^*_{h,2};\,\Theta_h)}{\mathcal{W}_{i,1}\mathcal{L}^F(\log w^*_{h,2};\,\Theta_i)}=\infty,\ \forall x\ \textit{an element of}\ \Theta_h,$$
$$\lim_{x'\to\infty}\frac{\mathcal{W}_{i,1}\mathcal{L}^F(\log w^*_{h,2};\,\Theta_i)}{\mathcal{W}_{h,1}\mathcal{L}^H(\log w^*_{h,2};\,\Theta_h)}=\infty,\ \forall x'\ \textit{an element of}\ \Theta_i.$$

Proof. Equilibrium $w^*_{h,2}$ as a function of Θ_i and Θ_h is determine by

$$\mathcal{W}_{h,1}\mathcal{L}^H(\log w^*_{h,2};\,\Theta_h)+\mathcal{W}_{i,1}\mathcal{L}^F(\log w^*_{h,2};\,\Theta_i)=L_{h,2}.$$

Therefore, the *equilibrium* $\mathcal{W}_{h,1}\mathcal{L}^H(\log w^*_{h,2};\Theta_h)$, which takes into the endogenous response of $\log w^*_{h,2}$, increases in $x, \forall x$ an element of Θ_h.[9] Furthermore, note that $\lim_{w_{h,2}\to\infty}\mathcal{L}^F(\log w_{h,2};\Theta_i)=0$ and that $\forall x$ an element of Θ_h, $\lim_{x\to\infty} w^*_{h,2}=\infty$. We thus have $\forall x$ an element of Θ_h, $\lim_{x\to\infty}\frac{\mathcal{W}_{h,1}\mathcal{L}^H(\log w^*_{h,2};\,\Theta_h)}{\mathcal{W}_{i,1}\mathcal{L}^F(\log w^*_{h,2};\,\Theta_i)}=\infty$.

It follows from the same argument that the equilibrium $\mathcal{W}_{i,1}\mathcal{L}^F(\log w^*_{h,2};\Theta_i)$ increases in x', for x' an element of Θ_i and that $\lim_{x'\to\infty}\frac{\mathcal{W}_{i,1}\mathcal{L}^F(\log w^*_{h,2};\,\Theta_i)}{\mathcal{W}_{h,1}\mathcal{L}^H(\log w^*_{h,2};\,\Theta_h)}=\infty$. □

Proof of Proposition 3.

Proof. From Lemmas B.3 and B.4, the labor market clearing condition of country h at period 2 can be written as

$$\mathcal{W}_{h,1}\mathcal{L}^H(\log w_{h,2};\,\Theta_h)+\mathcal{W}_{i,1}\mathcal{L}^F(\log w_{h,2};\,\Theta_i)=L_{h,2}. \tag{B.21}$$

Take total differentiation of (B.21) with respect to $\bar{\eta}_{ih,1}$,

$$\frac{d\log w_{h,2}}{d\bar{\eta}_{ih,1}}=-\frac{\frac{d\log\mathcal{W}_{h,1}}{d\bar{\eta}_{ih,1}}\mathcal{W}_{h,1}\mathcal{L}^H(\log w_{h,2};\,\Theta_h)+\frac{d\log\mathcal{W}_{i,1}}{d\bar{\eta}_{ih,1}}\mathcal{W}_{i,1}\mathcal{L}^F(\log w_{h,2};\Theta_i)}{\mathcal{L}^H_1(\log w_{h,2};\,\Theta_h)\mathcal{W}_{h,1}+\mathcal{L}^F_1(\log w_{h,2};\,\Theta_i)\mathcal{W}_{i,1}},$$

in which $\mathcal{L}^H_1(\log w_{h,2};\,\Theta_h)\equiv\frac{\partial\mathcal{L}^H(\log w_{h,2};\,\Theta_h)}{\partial\log w_{h,2}}$ and $\mathcal{L}^F_1(\log w_{h,2};\,\Theta_i)\equiv\frac{\partial\mathcal{L}^F(\log w_{h,2};\,\Theta_i)}{\partial\log w_{h,2}}$.

From Lemmas B.3 and B.4, $\mathcal{L}^H_1(\log w_{h,2};\,\Theta_h)<0$, $\mathcal{L}^F_1(\log w_{h,2};\,\Theta_i)<0$, and from Lemma B.5, $\mathcal{W}_{h,1}$ and $\mathcal{W}_{i,1}$ are independent of Θ_h or Θ_i. Therefore, holding Θ_i constant, $\frac{d\log w_{h,2}}{d\bar{\eta}_{ih,1}}<0$ if and only if:

$$\frac{d\log\mathcal{W}_{h,1}}{d\bar{\eta}_{ih,1}}\mathcal{W}_{h,1}\mathcal{L}^H(\log w_{h,2};\,\Theta_h)+\frac{d\log\mathcal{W}_{i,1}}{d\bar{\eta}_{ih,1}}\mathcal{W}_{i,1}\mathcal{L}^F(\log w_{h,2};\,\Theta_i)>0$$

Further noting that $\frac{d\log\mathcal{W}_{h,1}}{d\bar{\eta}_{ih,1}}<0$, $\frac{d\log\mathcal{W}_{i,1}}{d\bar{\eta}_{ih,1}}>0$, then from Lemma B.6, for a given x an element of Θ_h, the inequality holds when x is sufficiently large; it is violated when x is sufficiently small. Specifically, the

[9] This differs from Lemma B.3 in that here we also incorporate the impact of the increase in x on equilibrium $\log w^*_{h,2}$. We prove by contradiction. If not, then $\mathcal{W}_{i,1}\Theta_i\mathcal{L}^F(\log w^*_{h,2})$ must increase, which implies that $\log w^*_{h,2}$ must decline as Θ_i stays the same. The decrease in $\log w^*_{h,2}$ in turn implies an increase in the equilibrium $\mathcal{W}_{h,1}\mathcal{L}^H(\log w^*_{h,2};\Theta_h)$, resulting in a contradiction.

threshold value, x^*, is determined by

$$\frac{d\log \mathcal{W}_{h,1}}{d\bar{\eta}_{ih,1}}\mathcal{W}_{h,1}\mathcal{L}^H(\log w_{h,2};\Theta_{h,-x},x^*) + \frac{d\log \mathcal{W}_{i,1}}{d\bar{\eta}_{ih,1}}\mathcal{W}_{i,1}\mathcal{L}^F(\log w_{h,2};\ \Theta_i) = 0,$$

where $\Theta_{h,-x}$ is the vector of all other elements of Θ_h excluding x.

Similarly, holding Θ_h constant, $\frac{d\log w_{h,2}}{d\bar{\eta}_{ih,1}} < 0$ if and only if Θ_i is such that

$$\frac{d\log \mathcal{W}_{h,1}}{d\bar{\eta}_{ih,1}}\mathcal{W}_{h,1}\mathcal{L}^H(\log w_{h,2};\ \Theta_h) + \frac{d\log \mathcal{W}_{i,1}}{d\bar{\eta}_{ih,1}}\mathcal{W}_{i,1}\mathcal{L}^F(\log w_{h,2};\ \Theta_i) > 0,$$

which holds for any x' an element of Θ_i, if and only if x' is sufficiently small. This proves part (i) of the Proposition.

For part (ii), abusing the notation and write the foreign labor demand at period 2 as $\mathcal{L}^F(\log w_{h,2};\Theta_{i,-\bar{\eta}},\bar{\eta}_{ih,2})$, where $\Theta_{i,-\bar{\eta}}$ is all elements of $\Theta_{i,-\bar{\eta}}$ excluding $\bar{\eta}_{ih,2}$, to highlight the dependence of $\mathcal{L}^F$ on $\bar{\eta}_{ih,2}$ (see $\mathcal{L}^F$ definition in Lemma B.4).

By the definition of $\mathcal{L}^F$, we have $\mathcal{L}^F_1(\log w_{h,2};\ \Theta_{i,-\bar{\eta}},\bar{\eta}_{ih,2}) \equiv \frac{\partial \mathcal{L}^F(\log w_{h,2};\ \Theta_i)}{\partial \log w_{h,2}} < 0$ and $\mathcal{L}^F_3(\log w_{h,2};\Theta_{i,-\bar{\eta}},\bar{\eta}_{ih,2}) \equiv \frac{\partial \mathcal{L}^F(\log w_{h,2};\Theta_{i,-\bar{\eta}},\bar{\eta}_{ih,2})}{\partial \bar{\eta}_{ih,2}} > 0$. Now consider $w_{h,2}$ under any given $(\bar{\eta}_{ih,1},\bar{\eta}_{ih,2})$ determined by

$$\mathcal{W}_{h,1}\mathcal{L}^H(\log w_{h,2};\ \Theta_h) + \mathcal{W}_{i,1}\mathcal{L}^F(\log w_{h,2};\Theta_{i,-\bar{\eta}},\bar{\eta}_{ih,2}) = L_{h,2}. \tag{B.22}$$

We consider $\frac{d\log w'_{h,2}}{d\Delta\bar{\eta}'_{ih}}$, with $\Delta\bar{\eta}'_{ih}$ being the common value raised for $\bar{\eta}_{ih,1}$ and $\bar{\eta}_{ih,2}$, i.e., $\bar{\eta}'_{ih,1} = \bar{\eta}_{ih,1} + \Delta\bar{\eta}'_{ih}$, $\bar{\eta}'_{ih,2} = \bar{\eta}_{ih,2} + \Delta\bar{\eta}'_{ih}$, and $w'_{h,2}$ determined by (B.22) at $(\bar{\eta}'_{ih,1},\bar{\eta}'_{ih,2})$. Take total differentiation of (B.22) with respect to $\Delta\bar{\eta}'_{ih}$ and evaluate it at $(\bar{\eta}'_{ih,1},\bar{\eta}'_{ih,2})$:

$$\frac{d\log w'_{h,2}}{d\Delta\bar{\eta}'_{ih}} = -\frac{\frac{d\log \mathcal{W}_{h,1}}{d\bar{\eta}'_{ih,1}}\mathcal{W}_{h,1}\mathcal{L}^H(\log w'_{h,2};\ \Theta_h) + \frac{d\log \mathcal{W}_{i,1}}{d\bar{\eta}'_{ih,1}}\mathcal{W}_{i,1}\mathcal{L}^F(\log w'_{h,2};\ \Theta_{i,-\bar{\eta}},\bar{\eta}'_{ih,2}) + \mathcal{L}^F_3(\log w'_{h,2};\Theta_{i,-\bar{\eta}},\bar{\eta}'_{ih,2})\mathcal{W}_{i,1}}{\mathcal{L}^H_1(\log w'_{h,2};\Theta_h)\mathcal{W}_{h,1} + \mathcal{L}^F_1(\log w'_{h,2};\Theta_{i,-\bar{\eta}},\bar{\eta}'_{ih,2})\mathcal{W}_{i,1}}.$$

Since $\mathcal{W}_{h,1}$ and $\mathcal{W}_{i,1}$ are independent of Θ_h or Θ_i, and $\mathcal{L}^H_1(\log w'_{h,2};\ \Theta_h) < 0$, $\mathcal{L}^F_1(\log w'_{h,2};\ \Theta_i) < 0$, holding $\Theta_{i,-\bar{\eta}}$ constant, $\frac{d\log w'_{h,2}}{d\Delta\bar{\eta}'_{ih}} < 0$ if and only if Θ_h satisfies

$$\frac{d\log \mathcal{W}_{h,1}}{d\bar{\eta}_{ih,1}}\mathcal{W}_{h,1}\mathcal{L}^H(\log w'_{h,2};\ \Theta_h) + \frac{d\log \mathcal{W}_{i,1}}{d\bar{\eta}_{ih,1}}\mathcal{W}_{i,1}\mathcal{L}^F(\log w'_{h,2};\ \Theta_{i,-\bar{\eta}},\bar{\eta}'_{ih,2}) + \mathcal{L}^F_3(\log w'_{h,2};\ \Theta_{i,-\bar{\eta}},\bar{\eta}'_{ih,2})\mathcal{W}_{i,1} > 0. \tag{B.23}$$

Focusing on an element x of Θ_h, by Lemma B.6, given the value of other parameters, there exits a threshold value $\bar{x}(\Delta\bar{\eta}'_{ih})$ which depends on the raised value $\Delta\bar{\eta}'_{ih}$ and the given other parameters, such that (B.23) holds if $x > \bar{x}(\Delta\bar{\eta}'_{ih})$.

Define $\bar{\bar{x}} \equiv \sup_{\Delta\bar{\eta}'_{ih}\in[0,\Delta\bar{\eta}_{ih}]}\bar{x}(\Delta\bar{\eta}'_{ih})$, then we have $\forall x > \bar{\bar{x}}$, $\frac{d\log w'_{h,2}}{d\Delta\bar{\eta}'_{ih}} < 0, \forall\Delta\bar{\eta}'_{ih} \in [0,\Delta\bar{\eta}_{ih}]$. Thus, we have the second period wage relative to the autarky benchmark given by

$$\log\hat{w}_{h,2} - \log\bar{w}_{h,2} = \int_0^{\Delta\bar{\eta}_{ih}}\frac{d\log w'_{h,2}}{d\Delta\bar{\eta}'_{ih}}d\Delta\bar{\eta}'_{ih} < 0.$$

By the same rationale, we can show that holding other parameters constant, we can choose small enough x' for x' an element of $\Theta_{i,-\bar{\eta}}$ so that (B.23) holds. □

B.3 A Dynamic Model of MP without FDI

In this section, we describe a dynamic model of MP without FDI, which is used for comparison in Section 5.3. This alternative no-FDI model deviates from the benchmark model in two aspects: (1) λ_i and μ_h are set to infinity in all countries, i.e., there exist no financial frictions; (2) each firm can only produce at most in one host country, and the production function is modified to

$$y = z_{ih}^{\varphi}(k^{\alpha}l^{1-\alpha})^{1-\varphi}, \varphi \in (0,1)$$

which ensures a well-defined firm boundary under a perfect capital market; φ will be calibrated so that the foreign profit share implied by the modified model agrees with the one in the benchmark, to be detailed below.

Under these deviations, firms' optimal scale of production does not depend on their net worth. Furthermore, under the maintained assumption that there is a global capital market with a common world interest rate, firms are indifferent between finance from internal versus external sources, host versus home countries. They supply their net worth in this global capital market and borrow as necessary for production. Capital flows take place out of the firm boundary. In this sense, this is a model without FDI, with cross-border technology and capital transfers detached from each other. From this discussion, it should also be clear that capital market imperfections are the key to the rise of FDI and the exact reason why technology and capital transfers within MNEs are generally inseparable.

Below we characterize firms' decisions and the evolution of aggregate states in this alternative model. Lemma B.7 characterizes affiliates' production and financing decisions, analogous to Lemma 1; Lemma B.8 characterizes parent firms' overseas production decisions and expected returns under the return shocks, analogous to Lemma 3; Lemma B.9 characterizes the evolution of aggregate states, analogous to equation (9) and its characterization in Appendix B.2.4.

Lemma B.7. *Under the modified setup, for a firm from home country i with productivity z, once deciding to enter host country h, its factor use $k_{ih}(z)$ and $l_{ih}(z)$, production $y_{ih}(z)$, and total returns $R_{ih}(z)$ are independent of its net worth and characterized by*

$$\begin{aligned} k_{ih}(z) &= \frac{\alpha\kappa_{ih}m_{ih}(\tilde{z}_{ih}(z))}{r_h^b + \delta} \\ l_{ih}(z) &= \frac{(1-\alpha)\kappa_{ih}m_{ih}(\tilde{z}_{ih}(z))}{w_h} \\ y_{ih}(z) &= \tilde{z}_{ih}(z)^{\varphi}(k_{ih}(z)^{\alpha}l_{ih}(z)^{1-\alpha})^{1-\varphi} \\ R_{ih}(z) &= y_{ih}(z) - (\delta + r_h^b)k_{ih}(z) - w_h l_{ih}(z), \end{aligned}$$

where

$$\kappa_{ih} \equiv \alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)}(r_h^b+\delta)^{\alpha} w_h^{1-\alpha}$$
$$m_{ih}(z_{ih}) \equiv \Big(\frac{1-\varphi}{\kappa_{ih}}\Big)^{\frac{1}{\varphi}} z_{ih}.$$

Proof. With the integrated bond market and no financial constraints, the problem of the affiliate becomes

$$\tilde{R}_{ih}(z, e_h) = \max_{b_h^F, k, l, y} y + (1-\delta)k - w_h l - (1+r_h^b) b_h^F,$$
$$s.t. \quad y = \tilde{z}_{ih}(z)^{\varphi}(k^{\alpha} l^{1-\alpha})^{1-\varphi},$$
$$0 \le k \le e_h + b_h^F,$$

which is the same as the affiliate's problem in the benchmark model (equation (3)) except removing the constraint on b_h^F. The policies for k, l, y are then obtained by the firs-order conditions and are independent of e_h. Since e_h earns the same return at the host or the home country ($r_b^h = r_i^b$ under the integrated global bond market specification), the allocation of a firm's own net worth is undetermined. Without loss of generality we assume $e_h = 0, \forall h$ so the funds are supplied by the parent firm directly. This gives the return function in the lemma. □

Lemma B.8. *Under the modified setup, the fraction of firms from home country i with productivity z that operates in host country h is independent of firms' net worth. Denote this fraction $\chi_{ih}(z)$; it is characterized by:*

$$\chi_{ih}(z) = \Big(\frac{\tilde{R}_i(z)}{\overline{R}_i(z)}\Big)^{\theta} \cdot \frac{1}{N}\Big(\frac{\bar{\eta}_{ih} R_{ih}(z)}{\tilde{R}_i(z)}\Big)^{\theta}, \text{ for } h \quad \text{s.t. } \bar{\eta}_{ih} R_{ih}(z) < \overline{R}_i(z)$$
$$\chi_{ih}(z) = \Big[1 - \Big(\frac{\tilde{R}_i(z)}{\overline{R}_i(z)}\Big)^{\theta}\Big] + \Big(\frac{\tilde{R}_i(z)}{\overline{R}_i(z)}\Big)^{\theta} \cdot \frac{1}{N}\Big(\frac{\bar{\eta}_{ih} R_{ih}(z)}{\tilde{R}_i(z)}\Big)^{\theta}, \text{ for } h \text{ that attains } \bar{\eta}_{ih} R_{ih}(z) = \overline{R}_i(z)$$

where $\overline{R}_i(z) \equiv \max_{h'} \bar{\eta}_{ih'} R_{ih'}(z)$, $\tilde{R}_{ih}(z) \equiv \Big(\frac{1}{N}\sum_{h'}[\bar{\eta}_{ih'} R_{ih'}(z)]^{\theta}\Big)^{\frac{1}{\theta}}$, and $R_{ih}(z)$ characterized in Lemma B.7. Furthermore, the expected total returns of these firms before realizations of the return shocks are given by

$$\mathbb{E}[R_i(z, \boldsymbol{\eta})|z] = \Big(1 - [\tilde{R}_i(z)/\overline{R}_i(z)]^{\theta}\Big)\overline{R}_i(z) + [\tilde{R}_i(z)/\overline{R}_i(z)]^{\theta}\Big(\frac{\theta}{\theta-1}\overline{R}_i(z)\Big).$$

Proof. Since a firm can only operate at most in one host country, a firm with productivity z and net worth a solves the problem to maximize its end-of-period wealth, $\mathcal{A}_i(z, a)$, which consists of returns from production domestically or overseas and the interests earned from its net worth:

$$\mathcal{A}_i(z, a) = ([\max_h R_{ih}(z)\eta_{ih}] + (1+r_i^b)a,$$

where $R_{ih}(z)$ is the return that is characterized in Lemma B.7. Under the maintained assumption that $\eta_{ih} = \bar{\eta}_{ih}\zeta_h$ and $(\zeta_h)_{h=1}^N$ follows the distribution that is stated in Assumption 2, we arrive at the characterization of $\chi_{ih}(z)$ stated in the lemma. □

Lemma B.9. *Under the modified setup and the assumption that firms save β fraction of the after-return wealth* *every period,*[10] *the aggregate states of the economy collapse to the total net worths held by firms in each country* $(W_i)_{i=1}^N$*, with the transition of current total net worths* W_i *to future total net worths* W_i' *characterized by*

$$W_i' = \beta\Big(\underbrace{W_i(1+r_i^b)}_{\text{returns from capital}} + \underbrace{\int_0^\infty \mathbb{E}[R_i(z,\boldsymbol{\eta})|z]\bar{f}_i(z)\,\mathrm{d}z\Big)}_{\text{returns from production}},$$

with $\mathbb{E}[R_i(z,\boldsymbol{\eta})|z]$ *characterized in Lemma B.8 and* $\bar{f}_i(z)$ *the stationary density of firms with productivity* z*.*

Proof. Imposing the assumption that firms save β fraction of the after-return wealth, and integrating the firm-level after-return wealth characterized in Lemma B.8 over z and $\boldsymbol{\eta}$ we have,

$$\begin{aligned} W_{i,t+1} = \int_0^1 a_{t+1}(\omega)d\omega &= \int_0^1 \beta[a_t(\omega)(1+r_{i,t}^b) + \max_h R_{ih,t}(\omega)\eta_{ih}]d\omega \\ &= \int \beta[a(1+r_{i,t}^b) + \max_h R_{ih,t}(z)\eta_{ih}]\Phi_{i,t}(a,z)G(d\boldsymbol{\eta})dadz \\ &= \beta\Big(W_{i,t}(1+r_{i,t}^b) + \int_0^\infty \mathbb{E}[R_{i,t}(z,\boldsymbol{\eta})|z]\bar{f}_{i,t}(z)dz\Big), \end{aligned}$$

where the first line integrates over firms' net worth at $t+1$ with $\omega \in [0,1]$ denoting an individual firm; the second line applies the Law of Large Number and writes the integration with the density functions; and the last line applies the definition of $W_{i,t}$, $\mathbb{E}[R_{i,t}(z,\boldsymbol{\eta})|z]$ and $\bar{f}_i(z)$. □

Calibration of the alternative model. For calibration of the alternative model, first, λ_i, μ_h are set to infinity as discussed above. φ corresponds to the share of revenue earned as foreign affiliates' profits which is set to 0.366, the average foreign profit share in the benchmark model. Setting foreign profit shares equalized across models serves the purpose of isolating the dynamic effects of FDI via profit shifting and firm reinvesting that are emphasized in Section 5.3. We recalibrate $\alpha = 0.133$ so the alternative model implies the same static wage gains-MP share elasticity as in the benchmark model, which implies the two models, when calibrated to the same bilateral MP, delivers the same *static* wage gains (see Lemma B.10 below). Finally we assume that firms exogenously save β fraction of after-return wealth, which ensures the evolution of aggregate net worths comparable to the benchmark.

Following a similar procedure as in the benchmark, we set the times series of $\bar{\eta}_{ih}$ and $\bar{z}_i$ so that the following model-implied time series agree with the data counterparts: capital/GDP ratios, GDP per unit of labor of each country; and that bilateral multinational production as shares of receiving countries' total production agrees with the benchmark.

With the above calibration procedure, the only difference between the two models is that in the alternative model, due to the absence of financial constraints, firms' net worth (or in other words, accumulated retained earnings) no longer affects the current MP. It turns out that this difference is important for evaluating the dynamic gains from MP.

Lemma B.10. *Under the modified setup, assume that the outward FDI from country h is restricted and the interest rate is held constant, then the contemporaneous change in workers' wage in country h in response to a change in*

[10] Due to the non-homogeneity introduced by the decreasing-return-to-scale production function, firms' policy function for the saving rate under the CRRA utility function is not wealth-independent any more. The restriction ensures that the difference between the benchmark and the alternative models arise from the FDI-MP connection, rather than from the differences in capital accumulation decisions.

inward FDI policy is:

$$\Delta\log(w_h) = -\frac{1}{1+(1-\alpha)(\frac{1}{\varphi}-1)}\Delta\log\left(\frac{Y_{hh}}{Y_h}\right), \tag{B.24}$$

where $\frac{Y_{hh}}{Y_h}$ is the share of production conducted by domestic firms.

Proof. Starting from

$$\begin{aligned} l_{ih}(z) &= \frac{(1-\alpha)\kappa_{ih}m_{ih}(\tilde{z}_{ih}(z))}{w_h} \\ &\Rightarrow l_{ih} \propto \kappa_{ih}m_{ih}w_h^{-1} \\ &\propto \kappa_{ih}^{1-\frac{1}{\varphi}}z_{ih}w_h^{-1} \\ &\propto z_{ih}(r_h^b+\delta)^{\alpha(1-\frac{1}{\varphi})}w_h^{(1-\alpha)(1-\frac{1}{\varphi})-1} \end{aligned}$$

Similarly,

$$y_{ih} \propto z_{ih}(r_h^b+\delta)^{\alpha(1-\frac{1}{\varphi})}w_h^{(1-\alpha)(1-\frac{1}{\varphi})}$$

Therefore, rewrite the labor market clearing condition as

$$\frac{\int z\phi_h(z)dz}{Y_{hh}/Y_h}(r_h^b+\delta)^{\alpha(1-\frac{1}{\varphi})}w_h^{(1-\alpha)(1-\frac{1}{\varphi})-1} = L_h,$$

and therefore,

$$\Delta\log(w_h) = -\frac{1}{1+(1-\alpha)(\frac{1}{\varphi}-1)}\Delta\log\left(\frac{Y_{hh}}{Y_h}\right), \text{ with } \Delta\log(r_h^b+\delta)=0.$$

□

B.4 Micro-Foundations

In B.4.1 and B.4.2, we present our preferred micro-foundations for the capital market imperfections that distort affiliates' and parents' decisions. In B.4.3, we then present an alternative micro-foundation for the affiliate problem that speaks to the fact that FDI often takes place via M&A.

B.4.1 The Affiliate's Problem: Moral Hazard

We consider the problem of an affiliate receiving e_h investment from the parent. Define

$$\pi_h(z_{ih}) \equiv \alpha z_{ih}\left(\frac{1-\alpha}{w_h}\right)^{(1-\alpha)/\alpha} + 1 - \delta.$$

As described in Lemma 1, $\pi_h(z_{ih})$ is then the operating profit corresponding to *one-unit* of capital under the optimal choice of labor. The firm can raise capital from a local partner, whose opportunity cost of capital is the risk-free return r_h^b. Let the size of the project be k, so the local partner needs to put up the balance, $k - e_h$.

The parent adapts its technology to the local environment in h. The adaptation process is uncertain. The probability of success is $\underline{p} \in (0, 1)$, in which case the return is $\pi(z_{ih}) \cdot k$; the probability of failure is $(1 - \underline{p})$, in which case the return is $\iota \cdot \pi(z_{ih}) \cdot k, \;\; \iota < 1$. The management can make an effort to increase the success rate to 100%. Alternatively, they can put effort into generating a intangible, non-transferable benefit that accrues to the parent firm but does not generates cash flows.[11]

Let this intangible value be $f(z_{ih}, \omega_h) \cdot k$. It depends on three factors. First, the productivity of the technology—firms with better technology, or management practices, can generate larger intangible returns as well. Second, a host specific factor ω_h, which captures how difficult it is for the management to divert effort/resources and can be loosely interpreted as contract enforcement in h. We assume $\frac{\partial f}{\partial \omega_h} < 0$, i.e., stricter contract enforcement (large ω_h) means lower return from diverting the resources. Third, it also depends on the scale of the project. As the size of the planned project increases, the return from this alternative use of resource also increases. We assume that $f(z_{ih}, \omega_h) < (1 - \underline{p})(1 - \iota)\pi_h(z_{ih})$, which implies that it is jointly efficient for the parent to exert effort.

The parent offers a contract to local investors, which specifies the size of the project, the investment from the local partner, and the return to each party in different scenarios. Because parent efforts are non-verifiable, the contract can only condition on the outcome of the adaption. Formally, the optimal effort-inducing contract solves the following problem:

$$
\begin{aligned}
&\max_{k, x_S, x_F} \pi(z_{ih})k - x_S, \\
&s.t. \quad x_S \geq (1 + r_h^b)(k - e_h), \text{ (PC)} \\
&(\pi(z_{ih})k - x_S) \geq \underline{p}(\pi(z_{ih})k - x_S) + (1 - \underline{p})(\iota \cdot \pi(z_{ih})k - x_F) + f(z, \omega_h) \cdot k, \text{ (IC)} \\
&\iota \cdot \pi(z_{ih})k - x_F \geq 0, \\
&\pi(z_{ih})k - x_S \geq 0.
\end{aligned}
$$

Given the investment from the parent e_h, the contract chooses the size of the project and the payout to local investors under success or failure outcomes (x_S and x_F, respectively).

The objective function is the return to the parent if it exerts effort, in which case the success rate is one. The first constraint is the participation constraint of local investors; the second is the incentive compatibility constraint faced by the parent; the last two are the constraints arising from limited liability. The solution to the contractual problem is characterized by the following proposition.

[11] Examples of such non-transferable benefits include tacit knowledge about the host and the brand recognition of the parent firm. The parent might divert affiliate resources for experimentation of new products, which generates knowledge useful to the parent but of little value to the local investor; alternatively, the parent can divert resources into marketing, which increase the overall brand recognition of the parent, but this recognition is not transferable to the local partner.

Proposition B.1. *In the optimal solution, the first three constraints bind, and the solutions are*

$$k^* = \frac{\Delta p(1+r_h^b)}{\Delta p(1+r_h^b) - \Big(\Delta p \pi(z_{ih}) - f(z_{ih}, \omega_h)\Big)} e_h = \frac{(1+r_h^b)}{(1+r_h^b) - \Big(\pi(z_{ih}) - \frac{f(z_{ih},\omega_h)}{\Delta p}\Big)} e_h$$

$$x_F^* = \iota \cdot \pi(z_{ih}) \cdot k^*,$$

$$x_S^* = (1+r_h^b)(k^* - e_h),$$

where $\Delta p = 1 - \underline{p}$.

Under the technical assumption that $(\pi(z_{ih}) - \frac{f(z_{ih},\omega_h)}{\Delta p}) < 1 + r_h^b$ holds, which means the non-effort return cannot be so low that the moral hazard problem is not binding, the above solution implies that first, there exists a constraint akin to a collateral constraint, forcing the leverage to be below $\frac{(1+r_h^b)}{(1+r_h^b) - \big(\pi(z_{ih}) - \frac{f(z_{ih},\omega_h)}{\Delta p}\big)}$; second, the local partner's return to investment is $\frac{x_S^*}{k^* - e_h} = 1 + r_h^b$.

Intuitively, when the adaption fails, everyone can infer that the management did not put effort in the right task. However, since affiliates are a limited liability entity separated from the parent, the maximum amount the local investor can receive in this state is $\iota \cdot \pi(z_{ih}) \cdot k^*$. This in turn implies that for the local investor to break even, the payment under success, x_S, needs to be large enough. A large x_S blunts the incentive for the parent to exert effort, and more so when $\frac{k}{e_h}$ is higher.[12] Thus, the solution characterize the highest leverage in which effort can be incentive compatible.

Note that the maximum local finance ratio increases in $(\pi_h(z_{ih}) - \frac{f(z_{ih},\omega_h)}{\Delta p})$. This captures the net effect of the effort on the maximum leverage. All else equal, if ω_h is lower, the feasible leverage will be lower—with poorer contract enforcement institutions, the return from diverting effort to the intangible production is higher and it thus becomes more difficult to motivate effort, limiting the leverage. Now consider the special case of $f(z_{ih}, \omega_h) = \Delta p(\pi_h(z_{ih}) - \omega_h)$,[13] under which we have $k^* = \frac{1+r_h^b}{(1+r_h^b) - \omega_h} e$.

Define $\mu_h \equiv \frac{1+r_h^b}{(1+r_h^b) - \omega_h}$, it makes clear the source of variation in the affiliate financing decision across countries and over time. Cross-sectionally, ω_h matters more; over time, heightened interest spread, due to the tightening of the credit market, discourage local investors from singing on board. Our benchmark model corresponds to this special case with μ_h treated as a reduced-form parameter.

In this model, the payoff to local investors resemble both equity and debt. It resembles equity, as it is determined by a state-contingent payout policy. It also resembles debt, as only one state (successful adaption) will realize in equilibrium, so local investors are earning a fixed rate of return which is exactly equal to the opportunity cost of their capital—the interest rate.

B.4.2 The Parent's Problem: Defaultable Bonds

This section develops a model of defaultable bonds for the parent that micro-found the collateral constraint at the parent level, specified in problem (4) of the text.

[12] In our setting, both limited liability and unobserved effort play a role. Without limited liability, the local partner can induce effort by making parents pay more in the event of a failure. Without unobserved effort, the contract can condition on effort and the optimal contract will feature a higher maximum leverage.

[13] The specialization essentially assumes that the extra gains from diverting efforts to produce the intangible led by a higher z_{ih} is just balanced off by the gains from using the efforts to produce the contracted output, so the moral hazard problem is not worsened by a higher or lower z_{ih}.

A parent firm in country i with net worth a determines whether and how to scale up, after the realization of productivity z and idiosyncratic wedge draws $\{\boldsymbol{\eta}_h\}_h$. It faces a *schedule* of equilibrium-determined bond price $q_i(z,a,b)$, in which b is the size of the bond to be repaid next period, and $q_i(z,a,b)$ is the value of the bond in the current period (i.e., $1/q_i(z,a,b)-1$ is the net interest rate for this bond, and $q_i(z,a,b)b$ is the cash received in the current period.) Bond prices are a function of (z,a,b) because firms' default decision depends and only depends on these variables, to be detailed below.

The total size of funds used for investment is thus $a+q_i(z,a,b)b$. The firm can default on the bond at the end of current period; the cost of default is that $\gamma \in (0,1)$ fraction of the firm's total funds will be confiscated. Defaulting also incurs a deadweight cost in the sense that only $\iota \in (0,1)$ fraction of the confiscated funds is directed to the lender. The firm strictly prefers to default on the bond if and only if the total benefit of defaulting is greater than the total benefit of paying back the bond:

$$\begin{aligned} R_i^a(z,\boldsymbol{\eta})[a+q_i(z,a,b)b]-\gamma(a+q_i(z,a,b)b) &> R_i^a(z,\boldsymbol{\eta})[a+q_i(z,a,b)b]-b \\ \Leftrightarrow b &> \frac{\gamma}{1-\gamma q_i(z,a,b)}a. \end{aligned} \tag{B.25}$$

The bond price schedule is determined by the zero-profit condition of the competitive financial intermediary, whose cost of capital is $1+r_i^b$:

$$\underbrace{\mathbb{1}\left(b \leq \frac{\gamma}{1-\gamma q_i(z,a,b)}a\right)\cdot b}_{\text{payment upon no default}} + \underbrace{\mathbb{1}\left(b > \frac{\gamma}{1-\gamma q_i(z,a,b)}a\right)\cdot \iota \cdot \gamma \cdot (a+q_i(z,a,b)b)}_{\text{confiscated funds upon default}} = q_i(z,a,b)b(1+r_i^b),$$

which gives the equilibrium bond price schedule that reads

$$q_i(z,a,b) = \begin{cases} \dfrac{1}{1+r_i^b}, & \text{if } \dfrac{b}{a} \leq \dfrac{\gamma}{1-\gamma/(1+r_i^b)}, \\ \dfrac{\iota\gamma}{1+r_i^b-\iota\gamma}\dfrac{a}{b}, & \text{if } \dfrac{b}{a} > \dfrac{\gamma}{1-\gamma/(1+r_i^b)}. \end{cases}$$

Given the price schedule, the firm's returns net of interest payments as a function of b, taking into account the default option reads

$$\mathcal{R}_i(z,a,b,\boldsymbol{\eta}) = \begin{cases} R_i^a(z,\boldsymbol{\eta})(a+\dfrac{1}{1+r_i^b}b)-b, & \text{if } \dfrac{b}{a} \leq \dfrac{\gamma}{1-\gamma/(1+r_i^b)} \\ (R_i^a(z,\boldsymbol{\eta})-\gamma)(a+q_i(z,a,b)b), & \text{if } \dfrac{b}{a} > \dfrac{\gamma}{1-\gamma/(1+r_i^b)} \end{cases}$$

With $\iota < 1$, the optimal choice in the non-default range of $\frac{b}{a}$ (i.e., $\frac{b}{a} \leq \frac{\gamma}{1-\gamma/(1+r_i^b)}$) offers a strictly higher return to the firm than in the default range, and $\mathcal{R}_i(z,a,b,\boldsymbol{\eta})$ is increasing in b in the non-default range.

Hence, the firm will always choose the maximum non-default leverage, given by

$$\begin{aligned} b_i^* &= \frac{\gamma}{1-\gamma/(1+r_i^b)}a \\ \Leftrightarrow \qquad \underbrace{q_i(z,a,b_i^*)b_i^*}_{\text{cash received in the current period}} &= \frac{\gamma}{1-\gamma+r_i^b}a. \end{aligned}$$

Interpreting this solution through the lens of the collateral constraint model presented in the main text, it implies a collateral constraint given by $b^H \leq \lambda_h a$, with $\lambda_h \equiv \frac{\gamma}{1-\gamma+r_i^b}$. This constraint is tighter when γ is smaller (poor institutions) or when r_i^b is higher (tighter credit market).

The main message from this model is that because defaulting is socially inefficient (captured by $\iota < 1$), the equilibrium bond price will be such that the firm chooses not to borrow beyond a non-defaulting threshold of the debt level; below that threshold, firms endogenously choose not to default, so the bond price will be the same as the risk free rate in equilibrium. This micro-founds the collateral constraint presented in Section 3.3 of the text.

B.4.3 Alternative Formulation of Affiliate Problem as M&A

This subsection presents a tractable alternative model of FDI in which the foreign parent chooses host investors' extent of involvement in the affiliate and then splits the profit with them via Nash bargaining. We show that this model nests the benchmark model as a special case where the choice of host investor involvement is limited and host investors' bargaining power approaches zero. This demonstrates how the benchmark model can also capture the essential features of FDI that takes place via M&A.

Supply of local investment. As in the benchmark model, at the beginning of each period, firms decide whether to be active and if so, where to invest. Active firms pick the host country that offers the highest return. Idle firms invest all their funds through a representative domestic mutual fund and earn an average return of r_h^M (instead of lending at the market interest rate as in the baseline model). The supply of funds to the representative mutual fund of country h is

$$B_h^S = \int \hat{b}_h^H(z,\boldsymbol{\eta})\mathbb{1}(\hat{b}_h^H(z,\boldsymbol{\eta}) < 0)\phi_h(z)dG(\boldsymbol{\eta})dz,$$

where $\hat{b}_h^H(z,\boldsymbol{\eta})$ is determined by a cutoff rule similar to that in Lemma 2, except that the return from being idle is r_h^M, to be characterized later, instead of r_h^b in the baseline model. The representative mutual fund allocates B_h^S first to affiliates of foreign firms. The remainder of B_h^S, after all such investment is made, will be sent to the bond market (to be lent to parents) to earn a risk-free rate. We view the affiliates receiving investment from this fund as joint ventures between the foreign parent and the domestic mutual fund. The mutual fund is an equity investor that splits the rent with the foreign parent. Moreover, it also involves in management and affects the productivity of the joint venture, as described below.

FDI mode of entry choice. Consider an active firm investing e_h in host h. It meets with the manager of the host mutual fund, chooses among a menu of different modes of cooperation for the joint venture (which determines affiliate finance and productivity). The two parties then bargain over the surplus for the proposed mode.

There are M modes to choose from. A 'mode' involves two components characterized by a couple

(μ_h^j, χ_h^j), $j = 0, 1, ..., M$. The finance component, μ_h^j, determines how much the local partner contribute as an equity partner. Under mode j, the host partner contributes μ_h^j for each dollar the foreign parent brings in. Following the notation in the text, we denote the policy for host investment $b_{ih}^{F,j}(z, e_h)$:

$$b_{ih}^{F,j}(z, e_h) = \mu_h^j \cdot e_h.$$

The equity share of the host partner is then $\frac{\mu_h^j}{1+\mu_h^j}$. The technology component, χ_h^j, captures the elasticity of affiliate productivity with respect to parent productivity:

$$\tilde{z}_{ih}(z, j) = z^{\chi_h^j} \bar{z}_h^{1-\chi_h^j}. \tag{B.26}$$

A higher χ_h^j means the productivity of an affiliate is more like the parent and less like a typical firm in the host country. When the local partner (the mutual fund) holds more shares, they likely have a larger influence over the production technology and management practice of the affiliate. To capture this, we assume that within each host h, μ_h^j and χ_h^j are negatively correlated across j—to use more host finance, the parent has to give away more influence on affiliate production; conversely, to benefit more from the host technology, the parent needs to give out more shares.[14] An affiliate can thus be viewed as a joint venture that combines technology and capital from both foreign parent and domestic partner, capturing the essence of M&A FDI. Without loss of generality, we assume that μ_h^j increases in j. A special case is when $j = 0$, which corresponds to $\mu_h^0 = 0$ and $\chi_h^0 = 1$. Under this choice, the affiliate depends solely on the parent for both finance and technology, so this resembles greenfield FDI that uses no host financing.

For any given choice of j for (μ_h^j, χ_h^j), the parent and the representative mutual fund bargain to split the surplus. The total operation profits of a joint venture, if formed under mode j, is

$$\pi_h(z^{\chi_h^j} z_h^{1-\chi_h^j}) \cdot (1 + \mu_h^j) \cdot e_h, \tag{B.27}$$

where $\pi_h(z) \equiv \alpha z \left(\frac{1-\alpha}{w_h}\right)^{(1-\alpha)/\alpha} + 1 - \delta$, is the operation profit for each unit of capital when labor is chosen optimally. Equation (B.27) makes it apparent that the central tradeoff in the mode choice is one between scale $(1 + \mu_h^j)$ and control $(\pi_h(z^{\chi_h^j} z_h^{1-\chi_h^j}))$.

The outside option of the parent is to operate the affiliate alone ($j = 0$), with the following return:

$$\pi_h(z) \cdot e_h,$$

with $\pi_h(z)$ defined as above. The outside option of the host mutual fund is to lend the pledged fund, $b_{ih}^{F,j}$, and earn a risk-free rate r_h^b. The total surplus from the joint venture is

$$\pi_h(z^{\chi_h^j} z_h^{1-\chi_h^j}) \cdot (1 + \mu_h^j) \cdot e_h - \pi_h(z) \cdot e_h - \mu_h^j (1 + r_h^b) \cdot e_h.$$

The surplus is split between the two parties via Nash bargaining. Assuming the bargaining power of the parent firm is ω and the bargaining power of the mutual fund is $1 - \omega$, then the return to the parent

[14] We take the negative profile as a reduced-form technological constraint that is specific to each country. Its micro-foundation is beyond the scope of this extension.

from investing e_h in host h under mode j is

$$\begin{aligned}\tilde{R}_{ih}(z, e_h, j) &\equiv e_h \cdot \Big(\pi_h(z) + \omega[\pi_h(z^{\chi_h^j}\bar{z}_h^{1-\chi_h^j})(1+\mu_h^j) - \pi_h(z) - \mu_h^j(1+r_h^b)]\Big) \\ &\equiv e_h \cdot R_{ih}(z, j). \end{aligned} \tag{B.28}$$

And the return under the optimal mode choice is:

$$\tilde{R}_{ih}(z, e_h) = e_h \cdot \max_j R_{ih}(z, j), \tag{B.29}$$

with the corresponding decision rule, $j_{ih}(z)$, defined as

$$j_{ih}(z) = \text{argmax}_{j'} R_{ih}(z, j'). \tag{B.30}$$

Empirical implications The aforementioned tradeoff between scale and control implies that the mode choice will vary across firms and host countries systematically. Consider first a highly productive parent firm entering a poor country (with low $\bar{z}_h$). Forming a joint venture might lead to such a large productivity loss that the parent prefers to carry out the project alone. On the other hand, when a less productive firm is looking to enter a high-income country, forming a joint venture offers two potential benefits: higher productivity and access to host finance.[15] These predictions appear consistent with the data (Nocke and Yeaple, 2008).

Closing the model Note that equations (B.28) and (B.29) are both static and linear in e_h. We can therefore solve the mode choice j as a function of productivity z only, as in equation (B.30). The solution generalizes the cutoff rule for $\hat{b}_{ih}^F(z)$ in the baseline model (describe in Lemma 1). Replacing $R_{ih}(z)$ defined in equation (5) of the baseline model with equation (B.29), then the rest of our model machinery follows, subject to the following adjustments on the capital market clearing condition and capital returns.

The aggregate demand from foreign affiliates for host equity investment is

$$B_h^F = \sum_i \int \hat{b}_{ih}^F(z)\psi_{ih}(z)dz,$$

where $\hat{b}_{ih}^F(z) = \mu_h^{j_{ih}(z)}$.

The return to the domestic mutual fund in host h from investing μ_h^j capital in a joint venture under mode j with a parent firm from country i with productivity z is

$$\pi_{ih}^e(z) \equiv (1-\omega)\Big(\pi_h(z^{\chi_h^{j_{ih}(z)}}\bar{z}_h^{1-\chi_h^{j_{ih}(z)}})(1+\mu_h^{j_{ih}(z)}) - \pi_h(z)\Big) + \omega(1+r_h^b)\mu_h^{j_{ih}(z)}.$$

Under an integrated world credit market, the *average* return from investing in the country h mutual

[15] The second benefit is smaller if the cost of capital in the host country is high.

fund r_h^M is determined by

$$\underbrace{B_h^S(1+r_h^M)}_{\text{Total return}} = \underbrace{(1+r_h^b)\cdot(B_h^S - B_h^F)\cdot \mathbb{1}(B_h^S \geq B_h^F)}_{\text{Debt return}} + \underbrace{[\mathbb{1}(B_h^F \geq B_h^S)\cdot \frac{B_h^S}{B_h^F} + \mathbb{1}(B_h^F < B_h^S)]\cdot \sum_i \int \pi_{ih}^e(z)\psi_{ih}(z)dz}_{\text{Equity return}}. \tag{B.31}$$

The left side of the equation is the total return to the mutual fund. The first term on the right side is the return from lending to domestic parents in the form of debt, of what is left from equity investment. r_h^b is common across countries h and denotes the world interest rate. The second term on the right is the total return from equity investment in all affiliates. When $B_h^F > B_h^S$, i.e., when the demand for equity from affiliates exceeds total supply of funds to the mutual fund, the equity investment will be rationed to $\frac{B_h^S}{B_h^F}$ fraction of the firms randomly.

Under country-specific credit market, the domestic credit market clearing condition is,

$$\begin{aligned} B_h^S &= B_h^F + B_h^H, \\ B_h^H &= \int \hat{b}_h^H(z)\mathbb{1}(\hat{b}_h^H(z) > 0)\phi_h(z)dz, \end{aligned} \tag{B.32}$$

that is, the total supply of capital to the mutual fund equals the sum of credit and equity demand. The domestic risk-free rate, r_h^b, clears the market and determines r_h^M analogously to equation (B.31).

The baseline model is a special case with the following restriction: $\omega = 1$; only two mode choices, $\mu_h^0 = 0$ and $\mu_h^1 > 0$; and one single productivity elasticity for both cases, $\chi_h^0 = \chi_h^1 > 0$. As this extension makes clear, the baseline model captures essential features associated with M&A. It also provides a justification for why host finance should be interpreted as capturing both equity and debt finance.

B.5 Isomorphism

The benchmark model adopts a homogeneous good assumption. In this subsection, we show that if capital is used as fixed costs for the production of differentiated varieties and if the fixed cost increases with real income, then an environment with CES aggregation and monopolistic competition is isomorphic to the benchmark setup.

Consider there being a single final good used for consumption and investment, which can be assembled using intermediate-goods varieties in any country h according to

$$Y_h = \left(\int_0^{M_h} q_h(\omega)^{\frac{\sigma-1}{\sigma}} d\omega\right)^{\frac{\sigma}{\sigma-1}}, \sigma > 1,$$

where M_h is the number of varieties in country h, and ω denote a differentiated variety. To establish isomorphism we assume there is no trade between countries, so M_h is the measure of varieties produced in country h.

Denote X_h the aggregate expenditure in country h, the demand function for each variety ω is thus

$$q_h(\omega) = \frac{X_h p_h(\omega)^{-\sigma}}{P_h^{1-\sigma}},$$

where P_h is the aggregate price index

$$P_h = \Big[\int_0^{M_h} p_h^{1-\sigma}(\omega)d\omega\Big]^{\frac{1}{1-\sigma}}.$$

Assume the fixed cost for opening each product line is $\left(\frac{X_h}{P_h}\right)^{\kappa}$, with κ being a positive parameter and $\frac{X_h}{P_h}$ being the real income of country h. This specification captures congestion from other firms' production and dampens a potential 'scale effect', and implies that entry cost rises with development (Bollard et al., 2016). After paying up the fixed cost, each product line produces a differentiated variety according to

$$y = z_{ih}^{\frac{1}{\sigma-1}} l,$$

in which $z_{ih}^{\frac{1}{\sigma-1}}$ is the affiliate productivity,[16] and production input labor l is recruited from the host country labor market at a competitive nominal wage rate $\tilde{w}_h$.

Each affiliate owns a collection of product lines, financed by both parents and local investors. After production, the affiliate can recover $1-\delta$ fraction of the setup cost for the product line, analogous to the non-depreciated capital in the neoclassical setup. For each *individual variety*, facing a downward sloping demand function derived from the CES preference, the affiliate solves:

$$\begin{aligned} \Pi(z_{ih}) &= \max_{p,q,l} pq - \tilde{w}_h l \\ s.t. \quad q &= \frac{X_h p^{-\sigma}}{P_h^{1-\sigma}} \\ q &= z_{ih}^{\frac{1}{\sigma-1}} l. \end{aligned}$$

The optimality condition gives

$$\begin{aligned} p &= \frac{\sigma}{\sigma-1}\frac{\tilde{w}_h}{z_{ih}^{\frac{1}{\sigma-1}}}, \\ q &= \frac{X_h\big(\frac{\sigma}{\sigma-1}\frac{\tilde{w}_h}{z_{ih}^{\frac{1}{\sigma-1}}}\big)^{-\sigma}}{P_h^{1-\sigma}}, \\ l &= \frac{X_h\big(\frac{\sigma}{\sigma-1}\tilde{w}_h\big)^{-\sigma}}{P_h^{1-\sigma}} z_{ih}, \\ \Pi(z_{ih}) &= \frac{1}{\sigma}\frac{X_h\big(\frac{\sigma}{\sigma-1}\tilde{w}_h\big)^{1-\sigma} z_{ih}}{P_h^{1-\sigma}}. \end{aligned}$$

The *rate* of return to capital investment in the product line (made in the form of final goods) is

$$\pi_h(z_{ih}) = \frac{\Pi(z_{ih})/P_h}{(X_h/P_h)^{\kappa}} + 1 - \delta = \frac{1}{\sigma}\Big(\frac{\sigma}{\sigma-1}\Big)^{1-\sigma}\left(\frac{\tilde{w}_h}{P_h}\right)^{1-\sigma}\left(\frac{X_h}{P_h}\right)^{1-\kappa} z_{ih} + 1 - \delta \tag{B.33}$$

[16] The elasticity of productivity in z_{ih} happening to be $\frac{1}{\sigma-1}$ can be viewed as a normalization of the scale of z_{ih}.

Now we move to characterizing the decisions of the parent. Assume each unit of equity investment by the parent in host h can be levered to open multiple product lines, and the maximum leverage ratio is $(1+\mu_h)$. The return to investment in host h, factoring in the leverage, is

$$\begin{aligned} R_{ih}(z) &= \max_{\hat{b}^F} \pi(\tilde{z}_{ih}(z))(1+\hat{b}^F) - (1+r_h^b)\hat{b}^F \\ s.t. &\quad 0 \leq \hat{b}^F \leq \mu_h. \end{aligned}$$

We can similarly derive the fraction of firms with productivity z from home country i investing in host country h, denoted by $\hat{e}_{ih}(z)$, and the return on net worth, $R_i^a(z,\boldsymbol{\eta})$, which accounts for the idiosyncratic return shocks as in Lemma 3 for the baseline model .

Denote $\psi_{ih}(z)$ the density of investments from country i to country h by parents with productivity z . At the optimal choice of the parent firm, $\psi_{ih}(z) = (1+\lambda_i)\hat{e}_{ih}(z)\phi_i(z)$ as in the benchmark model. Denote $\hat{k}_{ih}(z) = 1 + \hat{b}_{ih}^F(z)$ the measure of product lines for each unit of investment from the parent. The labor market clearing condition in host country h is

$$L_h = \left(\frac{X_h}{P_h}\right)^{1-\kappa} \sum_i \int_0^\infty \left(\frac{\sigma}{\sigma-1}\frac{\tilde{w}_h}{P_h}\right)^{-\sigma} \tilde{z}_{ih}(z)\hat{k}_{ih}(z)\psi_{ih}(z)dz. \tag{B.34}$$

The intermediate-goods market clearing condition states that

$$\sum_i \int_0^\infty \left(\frac{\sigma}{\sigma-1}\frac{\tilde{w}_h}{P_h}\right)^{1-\sigma} \tilde{z}_{ih}(z)\hat{k}_{ih}(z)\psi_{ih}(z)dz = \left(\frac{X_h}{P_h}\right)^{\kappa}. \tag{B.35}$$

A sequential equilibrium is a time sequence (time subscript omitted) of $\left(\frac{\tilde{w}_h}{P_h}, \frac{X_h}{P_h}, r_h^b, \phi_i(z)\right)$ s.t. the labor market, goods market and global bond market clear, and the transition of $\phi_i(z)$ is consistent with the policy functions and the Markov transition density of exogenous productivity processes.

Recall in the baseline model with homogeneous good, the unit return of investment is

$$\pi_h(z_{ih}) = \alpha z_{ih}\left(\frac{1-\alpha}{w_h}\right)^{(1-\alpha)/\alpha} + 1 - \delta, \tag{B.36}$$

the labor market clearing condition is

$$L_h = \sum_i \int_0^\infty l(w_h)\tilde{z}_{ih}(z)\hat{k}_{ih}(z)\psi_{ih}(z)dz, \tag{B.37}$$

where $l(w_h) = \left(\frac{1-\alpha}{w_h}\right)^{1/\alpha}$, and the goods market clearing condition can be written as

$$Y_h = \sum_i \int_0^\infty y(w_h)\tilde{z}_{ih}(z)\hat{k}_{ih}(z)\psi_{ih}(z)dz, \tag{B.38}$$

where

$$y(w_h) = \left(\frac{1-\alpha}{w_h}\right)^{(1-\alpha)/\alpha}.$$

Comparing equations (B.33), (B.34) and (B.35) with (B.36), (B.37) and (B.38), we see that by setting $\frac{1}{\sigma} = \alpha$, $\frac{\tilde{w}_h}{P_h} = w_h$, $\kappa = 1$, and $\frac{X_h}{P_h} = Y_h$, the model with CES preference and monopolistic competition is equivalent to the model under neoclassical production for all *static* conditions. Given the decision rule for capital accumulation, it follows that the dynamic behaviors of the two models coincide, too.

The assumption that the fixed cost of setting up a product line is proportional to the real income of a country is crucial for the isomorphism because it exactly offsets the scale effect. More generally, when $0 < \kappa < 1$, larger economies offer a higher return for each product line and hence a higher return to capital investment, so the two models are not isomorphic. Nevertheless, the CES-monopolistic competition setting still retains all the tractability of the baseline model, and could be used to study the interaction between FDI, growth and trade.

B.6 Extensions

B.6.1 Firm-Level Switching Cost

The baseline model assumes that firms repatriate their profits and invest again in desired destinations every period. This may generate more switching among destinations compared to the data. This subsection shows that the model can be extended tractably to allow for the 'stickiness' in firm-level destination decisions. This extension would be useful for firm-level analyses or answering questions at the business cycle frequency. Given the focus of our quantitative exercises on the *medium-run* dynamics (our two decomposition exercises both focus on an episode of 6-7 years) of *aggregate* FDI, we choose to abstract from this extension in the main analysis.

We introduce option values in firms' investment decisions by assuming state-dependent FDI return wedges. Recall that $\bar{\eta}_{ih}$ is the systematic component of the return wedge.[17] In the main text, we assume these wedges to be the same for all firms from i investing in h. In this extension, we allow $\bar{\eta}_{ih}$ to be dependent on where a firm operated in the previous period, denoted by h_{-1}, as below:

$$\bar{\eta}_{ih_{-1},h} = \bar{\eta}^0_{ih}\Big[1 + \iota \cdot \mathbb{1}(h_{-1} = h)\Big]. \tag{B.39}$$

In equation (B.39), η^0_{ih} captures the return wedge for a firm from i opening up an affiliate in h, if in the previous period the firm was inactive or operating an affiliate *outside* country h. If, on the other hand, the firm had an affiliate in host h in the previous period, the mean return from investment would be higher by a factor of $\iota > 0$. The difference in return captures the cost of starting up a new affiliate in or switching to a new host, which could include investment in both physical capital and intangibles such as a supplier network and local knowhow.

The timing of the decision is the same as in the benchmark: the parent firm sees the realization of $\boldsymbol{\zeta}$ and decides whether/where to operate (h), and how much to reinvest (e). Let $h_{-1} = 1, ..., N$ denote the host country of an active firm in the previous period, and slightly abusing the notation, let $h_{-1} = 0$

[17] In quantification we specify $\bar{\eta}_{ih}(z) = \bar{\eta}_{ih} z^{\eta_z}$. The description here omits z and time subscript, as in the main text.

denote a firm that stayed inactive in the previous period. The Bellman equation for the parent firm is:[18]

$$\begin{aligned}
v_{ih_{-1}}(z,\boldsymbol{\zeta},a) &= \max_{h=0,1,\dots,N}\left\{v_{ih_{-1},h}(z,\boldsymbol{\zeta},a)\right\} \\
v_{ih_{-1},h}(z,\boldsymbol{\zeta},a) &= \max_{a'}\log\left(R^a_{ih}(z)\bar{\eta}_{ih_{-1},h}\zeta_h a - a'\right) + \beta\mathbb{E}[v_{ih}(z',\boldsymbol{\zeta}',a')|z],\ h=1,2,\dots,N \\
v_{ih_{-1},0}(z,\boldsymbol{\zeta},a) &= \max_{a'}\log\left((1+r^b_i)\zeta_0 a - a'\right) + \beta\mathbb{E}[v_{ih}(z',\boldsymbol{\zeta}',a')|z], \qquad \text{(B.40)}
\end{aligned}$$

in which $v_{ih_{-1}}$ denotes the value of a parent firm from home country i operating in country h_{-1} in the previous period. v_{i0} simply refers to the value of a previously idle firm. The parent chooses the destination h that offers the highest value $v_{ih_{-1},h}$. Depending on whether a firm chooses to be active in the current period, $v_{ih_{-1},h}$ takes different forms. For active firms investing in h, then the net return to investment is given by $R^a_{ih}(z)\bar{\eta}_{ih}(h_{-1})\zeta_h a$, where $R^a_{ih}(z)$, defined as in equation (7), is the return on net worth conditional on the destination choice h and summarizes the optimal financing and production decisions after the choice of h is made. For idle firms, the return is simply $(1+r^b_i)\zeta_0 a$, i.e., the product of risk free rate, an idiosyncratic shock for being inactive, ζ_0, and total net worth.

We solve the optimization problem by guess and verify. Suppose the value function is given by

$$v_{ih_{-1},h}(z,\boldsymbol{\zeta},a) = \hat{v}_{ih_{-1},h}(z,\boldsymbol{\zeta}) + \frac{1}{1-\beta}\log(a),\ h=0,1,\dots,N. \qquad \text{(B.41)}$$

That is, the value function conditional on the choice of h can be written as the sum of two components. The first is a log function of current net worth; the second is a function that depends on the firm's productivity, operation status of the previous period, and realization of the current idiosyncratic shocks, but is independent of net worth. Then by our guess,

$$v_{ih_{-1}}(z,\boldsymbol{\zeta},a) = \underbrace{\max_h\left\{\hat{v}_{ih_{-1},h}(z,\boldsymbol{\zeta})\right\}}_{\equiv\hat{v}_{ih_{-1}}(z,\boldsymbol{\zeta})} + \frac{1}{1-\beta}\log(a). \qquad \text{(B.42)}$$

We define $\hat{v}_{ih_{-1}}(z,\boldsymbol{\zeta}) \equiv \max_h\left\{\hat{v}_{ih_{-1},h}(z,\boldsymbol{\zeta})\right\}$. It captures the value of the optimal location choice h that is independent of net worth.

Substituting equations (B.41) and (B.42) into the right hand side of (B.40), we obtain:

$$\begin{aligned}
v_{ih_{-1},h}(z,\boldsymbol{\zeta},a) &= \max_{a'}\left\{\log\left(R^a_{ih}(z)\bar{\eta}_{ih}(h_{-1})\zeta_h a - a'\right) + \beta\mathbb{E}[\hat{v}_{ih}(z',\boldsymbol{\zeta}')|z] + \frac{\beta}{1-\beta}\log(a')\right\},\ h\neq 0 \\
v_{ih_{-1},0}(z,\boldsymbol{\zeta},a) &= \max_{a'}\left\{\log\left((1+r^b_i)\zeta_0 a - a'\right) + \beta\mathbb{E}[\hat{v}_{i0}(z',\boldsymbol{\zeta}')|z] + \frac{\beta}{1-\beta}\log(a')\right\}. \qquad \text{(B.43)}
\end{aligned}$$

Conditional on the choice of h, the first order condition over a' in each of these two cases give:

$$\begin{aligned}
a'_{ih_{-1},h}(z,\boldsymbol{\zeta},a) &= \beta\left[R^a_{ih}(z)\bar{\eta}_{ih}(h_{-1})\zeta_h\right]a,\ h\neq 0 \\
a'_{ih_{-1},0} &= \beta(1+r^b_i)\zeta_0 a. \qquad \text{(B.44)}
\end{aligned}$$

[18] To highlight the state-dependent feature of systematic return wedges, we make the primitive shocks $\boldsymbol{\zeta}$ an explicit argument of the value function and policy function, instead of $\boldsymbol{\eta}$ which have already incorporated the systematic component.

Now we characterize the choice of h (the probabilistic rule before the realization of $\boldsymbol{\zeta}$). Plugging equation (B.44) into equation (B.43), we obtain:

$$
\begin{aligned}
v_{ih_{-1},h}(z,\boldsymbol{\zeta},a) &= \underbrace{\log\Big(\beta[R^a_{ih}(z)\bar{\eta}_{ih}(h_{-1})]\Big) + \beta\mathbb{E}[\hat{v}_{ih}(z',\boldsymbol{\zeta}')|z]}_{\equiv V_{ih_{-1},h}(z)} + \log(\zeta_h) + \frac{1}{1-\beta}\log(a),\ h\neq 0 \\
v_{ih_{-1},0}(z,\boldsymbol{\zeta},a) &= \underbrace{\log\Big(\beta(1+r^b_i)\Big) + \beta\mathbb{E}[\hat{v}_{ih}(z',\boldsymbol{\zeta}')|z]}_{\equiv V_{ih_{-1},0}(z)} + \log(\zeta_0) + \frac{1}{1-\beta}\log(a)
\end{aligned}
\tag{B.45}
$$

For both cases, the sum of the first two components is defined to be $V_{ih_{-1},h}(z)$. It captures the value that is independent of $\boldsymbol{\zeta}$ and net worth, if destination h is chosen and the subsequent consumption-saving decision is at the optimal. The location choice from the first line of equation (B.40) then reduces to

$$
V_{ih_{-1}}(z,\boldsymbol{\zeta}) \equiv \max_{h=0,\dots,N}\Big[V_{ih_{-1},h}(z)+\log(\zeta_h)\Big]
$$

At this point, it is necessary to make distributional assumptions on ζ_h. We deviate from the distributional assumption made in the baseline model, and specify $\log(\zeta_h)$ to be i.i.d. across parent firms and host countries, drawn from a Gumbel distribution with scale parameter $1/\theta \in (0,1)$, i.e., with probability distribution function (PDF)

$$
f_{\log(\zeta)}(x) = \theta\exp\Big(-\theta x - \exp(-\theta x)\Big), \forall x\in\mathbb{R}.
$$

We can then derive the probability that a firm with productivity z chooses location h:

$$
\hat{e}_{ih_{-1},h}(z) = \frac{\exp\Big[\theta V_{ih_{-1},h}(z)\Big]}{\sum_{\tilde{h}=0}^{N}\exp\Big[\theta V_{ih_{-1},\tilde{h}}(z)\Big]},
\tag{B.46}
$$

and the expected value of $V_{ih_{-1}}(z,\boldsymbol{\zeta})$ *before* the realization of $\boldsymbol{\zeta}$ and the choice of h:

$$
\begin{aligned}
\bar{V}_{ih_{-1}}(z) &\equiv \mathbb{E}[V_{ih_{-1}}(z,\boldsymbol{\zeta})|z] \\
&= \frac{\bar{\gamma}}{\theta} + \frac{1}{\theta}\log\Big(\sum_{\tilde{h}=0}^{N}\exp\Big[\theta V_{ih_{-1},\tilde{h}}(z)\Big]\Big),
\end{aligned}
\tag{B.47}
$$

where $\bar{\gamma}$ is the Euler-Mascheroni constant. Combining equations (B.45) and (B.47) and noticing $\mathbb{E}[v_{ih}(z',\boldsymbol{\zeta}')|z] = \mathbb{E}[\bar{V}_{ih}(z')|z]$ by the Law of Iterated Expectation, we have the following functionals for $\{\bar{V}_{ih_{-1}}(z), V_{ih_{-1},h}(z)\}$:

$$
\begin{aligned}
\bar{V}_{ih_{-1}}(z) &= \frac{\bar{\gamma}}{\theta} + \frac{1}{\theta}\log\Big(\sum_{\tilde{h}=0}^{N}\exp\Big[\theta V_{ih_{-1},\tilde{h}}(z)\Big]\Big) \\
V_{ih_{-1},h}(z) &= \beta\mathbb{E}[\bar{V}_{ih}(z')|z] + \begin{cases}\log\Big(\beta[R^a_{ih}(z)\bar{\eta}_{ih}(h_{-1})]\Big), \text{ if } h=1,\dots,N \\ \log\Big(\beta(1+r^b_i)\Big), \text{ if } h=0.\end{cases}
\end{aligned}
\tag{B.48}
$$

The above system of equations can be solved for $V_{ih_{-1},h}(z)$ as a function of z via standard numerical methods, such as value function iterations. We can plug the solution into equation (B.42) to verify that the guess specified in equation (B.41) holds. The probabilistic rule for location choice is then given by equation (B.46).

Aggregation. The expected return conditional on choosing location h is given by:[19]

$$\begin{aligned}\bar{R}^a_{ih}(z, h_{-1}) &\equiv \mathbb{E}\left[R^a_{ih}(z)\bar{\eta}_{ih}(h_{-1})\zeta_h \middle| z, h_{-1}, \text{location } h \text{ attains maximum}\right] \\ &= R^a_{ih}(z)\bar{\eta}_{ih}(h_{-1})\left(\sum_{\tilde{h}=0}^{N} \exp\left[\theta(V_{ih_{-1},\tilde{h}}(z) - V_{ih_{-1},h}(z))\right]\right)^{\frac{1}{\theta}} \Gamma\left(1 - \frac{1}{\theta}\right), \end{aligned} \tag{B.49}$$

where $\Gamma(\cdot)$ is the gamma function.

Utilizing that the decision rules are linear in the parent firm's net worth, to compute the aggregate quantities in the model, it is sufficient to track the wealth density of firms at each productivity level z by home country i and location choice in the previous period h_{-1}, denoted by $\phi_{ih_{-1}}(z)$. At the stationary equilibrium $\phi_{ih_{-1}}(z)$ should satisfy the following transition:

$$\phi_{ih}(z') = \sum_{h_{-1}} \int_0^{\infty} \phi_{ih_{-1}}(z)\hat{e}_{ih_{-1},h}(z)\beta\bar{R}^a_{ih}(z, h_{-1})f_i(z'|z)dz,$$

where $\bar{R}^a_{ih}(z, h_{-1})$ is the expected return on net worth conditional on choosing h, characterized in equation (B.49), and $f_i(z'|z)$ is the conditional density function for the Markov productivity process in country i. Due to the stickiness of the parent firm's investment decision, the aggregate state includes the wealth densities of firms conditional on their last-period location choice as well as their predetermined home country, increasing the complexity of computation by a factor that is equal to the number of countries, but the separability between z and a remains and the algorithm for the baseline model applies.

B.6.2 CRRA Utility

This subsection shows that the feature that makes the model tractable, the separability of parent productivity z and net worth a in decisions, holds when entrepreneurs have a more general constant-relative-risk-aversion (CRRA) utility. To shorten notations, we omit the time subscript.

The Bellman equation of the parent is given by:

$$\begin{aligned} v_i(z, \boldsymbol{\eta}, a) = &\max_{c, a', \{e_h\}_{h=1}^N, b^H} u(c) + \beta\mathbb{E}\left[v_i(z', \boldsymbol{\eta}', a')|z\right] \\ &s.t. \quad \sum_h e_h = a + b^H \\ &\quad -a \le b^H \le \lambda_i \cdot a \\ &\quad c + a' = \sum_h \tilde{R}_{ih}(z, e_h)\eta_{ih}e_h - (1 + r^b_i)b^H, \end{aligned}$$

with $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$. Parameter $\sigma > 0$ governs the degree of relative risk aversion. Because the change in

[19] To show this, first notice that $\left(V_{ih_{-1},h}(z) + \log(\zeta_h) \middle| h \text{ attains maximum}\right)$ follows a Gumbel distribution with location parameter $\frac{1}{\theta}\log\left(\sum_{\tilde{h}=0}^{N} \exp\left[\theta V_{ih_{-1},\tilde{h}}(z)\right]\right)$. Then apply that the exponential of a Gumbel-distributed random variable follows the Frechet distribution to get the desired results.

static utility does not affect static profit maximization at the affiliate level, Lemma 1 holds and moreover, the parent will follow the same cutoff rule for leverage at the home country, as specified in equation (6). The dynamic programming problem therefore reduces to:

$$v(z,\boldsymbol{\eta},a) = \max_{a'} u\left(R^a(z,\boldsymbol{\eta})a - a'\right) + \beta\mathbb{E}[v(z',\boldsymbol{\eta}',a')|z], \tag{B.50}$$

where $R_i^a(z,\boldsymbol{\eta})$ is as defined in equation (7).

We proceed using a guess and verify strategy. Guess the value function takes the form

$$v(z,\boldsymbol{\eta},a) = \hat{v}(z,\boldsymbol{\eta})\frac{a^{1-\sigma}}{1-\sigma}.$$

Plug the guess into the right hand side of Bellman equation (B.50), we have

$$\max_{a'} u(R^a(z,\boldsymbol{\eta})a - a') + \beta\mathbb{E}[\hat{v}(z',\boldsymbol{\eta}')|z]\frac{[a']^{1-\sigma}}{1-\sigma}.$$

The first order condition with respect to a' gives

$$a' = \frac{R^a(z,\boldsymbol{\eta})}{1+(\beta\mathbb{E}[\hat{v}(z',\boldsymbol{\eta}')|z])^{-1/\sigma}} \cdot a, \tag{B.51}$$

and

$$c = [\beta\mathbb{E}[\hat{v}(z',\boldsymbol{\eta}')|z]]^{-1/\sigma} \cdot a'.$$

Note that in this setting, expectation about future environment enters the current capital accumulation decision. Plug equation (B.51) into the right hand side of the Bellman equation:

$$v(z,\boldsymbol{\eta},a) = \frac{1}{1-\sigma}\left\{\frac{(\beta\mathbb{E}[\hat{v}(z',\boldsymbol{\eta}')|z])^{-1/\sigma}}{1+(\beta\mathbb{E}[\hat{v}(z',\boldsymbol{\eta}')|z])^{-1/\sigma}}R^a(z,\boldsymbol{\eta})a\right\}^{1-\sigma} + \frac{\beta\mathbb{E}[\hat{v}(z',\boldsymbol{\eta}')|z]}{1-\sigma}\left\{\frac{R^a(z,\boldsymbol{\eta})a}{1+(\beta\mathbb{E}[\hat{v}(z',\boldsymbol{\eta}')|z])^{-1/\sigma}}\right\}^{1-\sigma}.$$

Comparing this with the guess, $v(z,\boldsymbol{\eta},a) = \hat{v}(z,\boldsymbol{\eta})\frac{a^{1-\sigma}}{1-\sigma}$, gives

$$\hat{v}(z,\boldsymbol{\eta}) = \left\{\frac{R^a(z,\boldsymbol{\eta})}{1+(\beta\mathbb{E}[\hat{v}(z',\boldsymbol{\eta}')|z])^{-1/\sigma}}\right\}^{1-\sigma}\left\{(\beta\mathbb{E}[\hat{v}(z',\boldsymbol{\eta}')|z])^{-(1-\sigma)/\sigma} + \beta\mathbb{E}[\hat{v}(z',\boldsymbol{\eta}')|z]\right\}.$$

Integrating over $\boldsymbol{\eta}$, utilizing that $R^a(z,\boldsymbol{\eta})$ has closed-form representations we have

$$\mathbb{E}[\hat{v}(z,\boldsymbol{\eta})|z] = \mathbb{E}[R^a(z,\boldsymbol{\eta})^{1-\sigma}|z]\left[1+(\beta\mathbb{E}[\hat{v}(z',\boldsymbol{\eta}')|z])^{-1/\sigma}\right]^{\sigma-1}\left[(\beta\mathbb{E}[\hat{v}(z',\boldsymbol{\eta}')|z])^{-(1-\sigma)/\sigma} + \beta\mathbb{E}[\hat{v}(z',\boldsymbol{\eta}')|z]\right]. \tag{B.52}$$

By the Law of Iterated Expectation

$$\mathbb{E}[\hat{v}(z',\boldsymbol{\eta}')|z] = \mathbb{E}\Big(\mathbb{E}[\hat{v}(z',\boldsymbol{\eta}')|z']\Big|z\Big).$$

Denote

$$\tilde{v}(z) = \mathbb{E}[\hat{v}(z, \boldsymbol{\eta})|z],$$

then equation (B.52) reduces to

$$\tilde{v}(z) = \mathbb{E}[R^a(z, \boldsymbol{\eta})^{1-\sigma}|z]\left[1 + (\beta\mathbb{E}[\tilde{v}(z')|z])^{-1/\sigma}\right]^{\sigma-1}\left[(\beta\mathbb{E}[\tilde{v}(z')|z])^{-(1-\sigma)/\sigma} + \beta\mathbb{E}[\tilde{v}(z')|z]\right], \tag{B.53}$$

which defines a functionals for $\tilde{v}(z)$, for which we can solve using the standard method (noticing $\mathbb{E}[R^a(z, \boldsymbol{\eta})^{1-\sigma}|z]$ can be derived in closed forms similar to Lemma 3). Plugging the resulting $\tilde{v}(z)$ into equation (B.51), we arrive at the policy function for internal capital accumulation a'.

C Quantification

Section C.1 describes the construction of the data used for quantification. Section C.2 details the calibration procedures. Section C.3 conducts an additional validation exercise and highlights the general equilibrium effects by comparing the predictions of our multi-country experiments to the predictions of counterfactual experiments carried out on individual countries.

C.1 Additional Sources of Data for Calibration

GDP, capital, and effective employment. The time series of GDP, capital stock, and effective employment used in calibration are all obtained from the PWT 9.0. The data series for GDP and capital stock are *cgdpo* and *ck*, respectively. Effective employment is calculated as the product of employment (*emp*) and human capital (*hc*).

Bilateral FDI. Our quantification requires bilateral FDI stocks over 2001-2012. We construct this dataset by combining the bilateral FDI *stocks* for 2001 from Ramondo et al. (2015) and a panel of bilateral FDI *flows* for 2002-2012 we newly assemble from the publications of UNCTAD. With these data, we measure the total FDI stock from country i to country h at the end of year t as:

$$\text{FDI Stock}_{ih,t} = (1-\delta)\text{FDI Stock}_{ih,t-1} + \text{FDI Flow}_{ih,t}, \tag{C.1}$$

in which δ is the depreciation rate for capital, obtained from the PWT. We take the following steps to check the quality and ensure the consistency of the data.

Sample country representativeness for the World FDI. A few economies outside our sample might carry disproportionate weights in world FDI (some of these have more inward FDI than our sample countries, but are excluded because they are not in the Ramondo et al. (2015) dataset). This raises the question of whether the aggregate FDI in our sample aligns well with aggregate world FDI. Figure C.1 shows that, although the sum of FDI flows among our sample countries is smaller than the sum across all countries, the two series show a similar trend.[20] Moreover, a simple adjustment can reduce the difference between the two measures by half. Specifically, countries outside the sample but nonetheless

[20] Aggregate FDI between all countries in the world is separately reported by the UNCTAD.

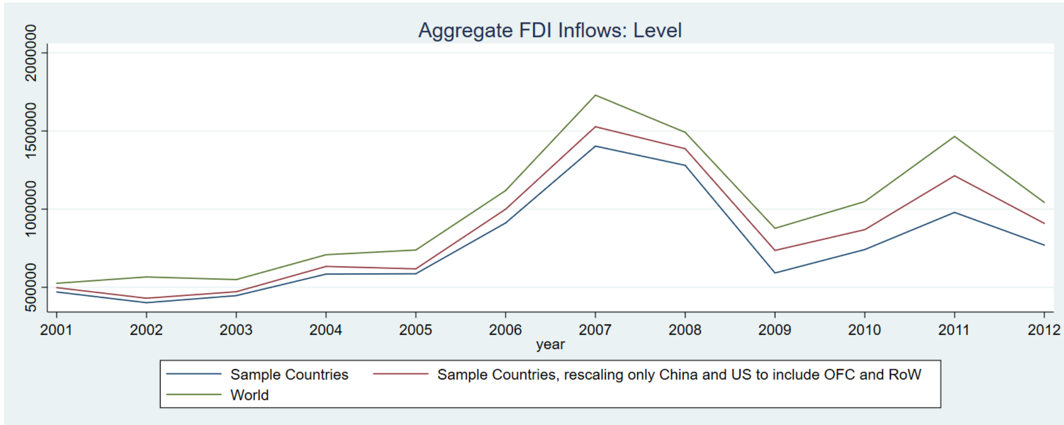

Figure C.1: Sample and World Aggregate FDI Inflow

Note: the figure shows the aggregate FDI between our sample countries (raw as well as adjusted for offshore financial centers for U.S. and China) and the aggregate world FDI. Units are in million current price USD.

attract/send a large amount of FDI are usually offshore financial centers, which act as intermediaries for investment in third countries. For example, Hong Kong has been an important intermediary for mainland Chinese enterprises to invest abroad and for the rest of the world to invest in China; similarly, as the most important sending country of FDI, the U.S. also invests heavily in Bahamas and Panama, presumably to direct investment elsewhere while avoiding taxes. Given that China (mainland) and the U.S. are among the largest countries, we add back the bilateral FDI involving U.S. and China that might have gone through the following offshore financial centers: Hong Kong, Cayman Islands, Bahamas, Dominica, British Virgin Island, Panama, and Luxembourg. We make a proportionate assumption in this adjustment by redistributing the FDI flows from China and the U.S. towards offshore financial centers to other sample countries based on the shares of these sample countries in outward FDI from China and the U.S., respectively. We make similar adjustments for inward FDI to China and the U.S. The red line in Figure C.1 shows that these adjustments alone reduce the gap by a half.

Missing values in the FDI data. Both the initial FDI stocks from Ramondo et al. (2015) and our subsequent flows have missing values (they are excluded in Figure C.1). Ramondo et al. (2015) impute stock values based on the reported number of affiliates. For quantification, we use this imputation (note that we do not use imputed values for empirical analysis). We assume the remaining missing values in the initial bilateral FDI stock are zeros.

The bilateral FDI flows in equation (C.1). For each pair of countries, we observe potentially two reported values for bilateral FDI flows, one from the host country and the other from the home country. When both values are non-missing, we use the number reported by the source country, as this likely captures the ultimate source country of investment more accurately. When both values are missing, we consider the following scenarios. First, if a country pair has a recorded zero in *all* years aside from the year with missing values, we replace each missing value with zero, too. Second, if a country pair has one or more gaps *between* a string of positive values, we impute these missing values. We calculate the shares of the destination country in the home country's total outward FDI flows in years with available data, linearly impute the shares, and calculate the corresponding bilateral flows based on the imputed shares and the home country's total outward FDI for the years with missing values. After these two procedures, the remaining missing values are mostly from country pairs with missing values for the entire sample period. To avoid extrapolation, we simply assume these values are zero.

Currency and price adjustments for FDI flows. The FDI flows from the UNCTAD are denominated in current price dollars, whereas the capital stock and GDP from the PWT are in constant prices. For consistency, we first calculate the ratios between FDI flows and GDP of destination countries in current price dollar in each year. We then multiply these ratios with the constant-price GDP from PWT to obtain the constant-price FDI flows.

Affiliate financing. We use the BEA data to construct the time series of affiliate financing compositions. Conceptually, an affiliate could be financed by investors of three different origins: the home country (including but not limited to the parent company), the host country, and third countries. Each of these three groups can then be further classified into creditors or equity investors. Given that our model is not designed to study capital structure choices, we will bundle equity and debt together. In our model, affiliate finance comes from only two sources: home and host countries. Presumably, most third-country investors only come on board because the backing of the parent firm. Through the lens of our model, therefore, it is most appropriate to treat this as funding from the parent country, raised with the parent net worth as collateral.

There are two separate measures for the share of finance from the host country. **The first measure** is based on the statistics on the composition of external finance, available for 1983-2008. During these years, the overseas affiliates of U.S. MNEs report their sources of finance from the three groups of countries (host, home, and others). This dataset can be used to directly construct the host finance share, but it is discontinued after 2009.

The second measure is constructed as the ratio of two variables: the total position of U.S. investment in affiliates in a host country, and the total assets of U.S. affiliates in a host country. These two series cover our entire sample period, but in these series third-country finance is grouped with host finance in reporting, which inflates the influence of host countries' financial markets.

We combine the two measures to construct the series of host country finance share. Between 1983 and 2008, we use the first measure. Between 2009 and 2012, we use the *initial level* from the first measure in 2008, together with the *yearly change* in percent of the second measure, to impute the host country finance share in the balance sheet. The underlying assumption for the imputation is that the yearly changes in these two measures after 2008 are correlated. This assumption cannot be tested directly, but we can test if it holds before 2008. Figure C.2 shows the residual plots when we regress one measure on the other, controlling for country and year fixed effects. It shows that indeed these two measures are correlated.

To construct an analogous measure for the U.S. as a host country, we apply a similar approach to the BEA data on the affiliates of foreign MNEs operating in the U.S.

There are two remaining challenges in constructing the series. First, between 1999 and 2003, the total value of host finance is missing from the first measure, but host *debt* finance is reported. In this case, we impute the total host finance by assuming that the share of host finance in all finance in 1999-2003 is equal to the share of host debt finance in all debt finance. As debt finance accounts for the majority of host finance, simply using the ratio of host debt finance in total finance yield similar trends.

Second, there is a change of definition in the second measure on whether it covers non-bank affiliates. In years in which statistics for both all affiliates and all non-bank affiliates are separately reported, we compare the host finance shares from these two samples. We find that they are very similar.

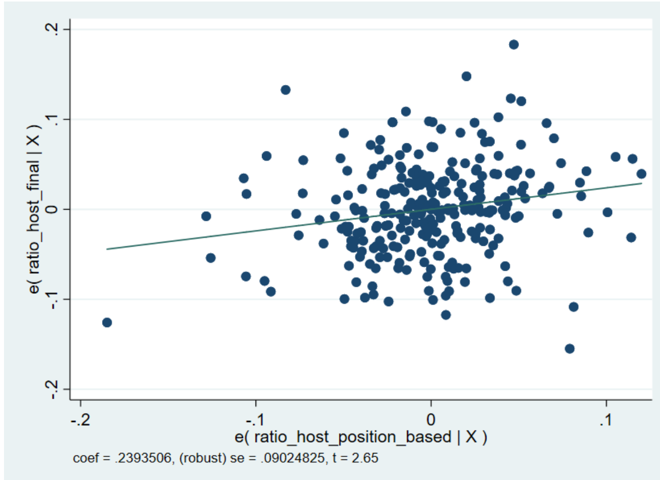

Figure C.2: Correlation Between Two Measures of Host Finance Share

Note: this figure shows that controlling for country and year fixed effects, the two measures of host finance share are significantly correlated with one another.

C.2 Calibration

C.2.1 Estimating η_z

We explain below how we estimate η_z in equation (14). In the model, when $\tilde{R}_i(z) = \overline{R}_i(z)$, the odds ratio of a firm becoming an MNE is

$$\frac{Pr(MN|z)}{1 - Pr(MN|z)} = \frac{\sum_{h' \neq i} \chi_{ih'}(z)}{\chi_{ii}(z)},$$

with $\chi_{ih}(z)$ defined in equation (B.12). The log odds ratio is therefore

$$\log\left(\frac{Pr(MN|z)}{1 - Pr(MN|z)}\right) = \log\left(\frac{\sum_{h' \neq i} [\bar{\eta}_{ih'} R_{ih'}(z)]^{\theta}}{[\bar{\eta}_{ii} R_{ii}(z)]^{\theta}}\right) + \theta \cdot \eta_z \log(z).$$

Generally, $\log(\frac{\sum_{h' \neq i} [\bar{\eta}_{ih'} R_{ih}(z)]^{\theta}}{[\bar{\eta}_{ii} R_{ii}(z)]^{\theta}})$ depends on parent's productivity z because affiliates' financing decision depends on it. However, for firms that are productive enough that their rent is much higher than the the interest rate, the following holds:

$$\lim_{z \to \infty} \frac{\sum_{h' \neq i} [\bar{\eta}_{ih'} R_{ih'}(z)]^{\theta}}{[\bar{\eta}_{ii} R_{ii}(z)]^{\theta}} = \frac{\sum_{h' \neq i} [\bar{\eta}_{ih'} \bar{z}_{h'}^{1-\gamma} w_{h'}^{\frac{\alpha-1}{\alpha}}]^{\theta}}{[\bar{\eta}_{ii} \bar{z}_{i}^{1-\gamma} w_{i}^{\frac{\alpha-1}{\alpha}}]^{\theta}}.$$

This ratio is i specific and does not depend on the firm-specific productivity z. Therefore, for these firms we have the following structural equation:

$$\log(\frac{Pr(MN|z)}{1 - Pr(MN|z)}) \approx \beta_{0i} + \theta \cdot \eta_z \log(z),$$

in which β_{0i} is the country fixed effect. This specification corresponds exactly to a Logit regression model, with the probability of a country-i firm being an MNE being

$$Pr(MN|i,z) = F(\beta_{0i} + \beta_1 \log(z)) = \frac{\exp(\beta_{0i} + \beta_1 \log(z))}{1 + \exp(\beta_{0i} + \beta_1 \log(z))}. \tag{C.2}$$

Because this is only an exact structural equation when $\tilde{R}_{i,t}(z) = \overline{R}_{i,t}(z)$ *or* for firms with large enough productivity, β_1 does not admit a structural interpretation. We adopt an indirect-inference procedure to estimate the structural parameter.

Specifically, we estimate equation (C.2) using the dataset described in Bloom et al. (2012), a representative survey of manufacturing firms covering a dozen of countries. Given that the structural equation is a good approximation only for firms that are productive, we keep only firms in top 25% of the productivity distribution in each country and regress a dummy of whether a firm is an MNE on the log of TFP, controlling for country fixed effects. We find a point estimate of 0.184 for $\hat{\beta}_1$. We then pick η_z so that when estimated using the model-simulated data that are from the common set of countries as in Bloom et al. (2012), equation (C.2) produces the same elasticity. This determines $\eta_z = 0.03$.[21]

C.2.2 Algorithm for Solving and Calibrating the Model

Solving the steady state

1. Choose a country-specific predetermined grid of size N_z for firm productivity z. The range of the grid covers 5 unconditional standard deviations of the country's productivity process. The grid is evenly spaced over the value of $\log(z)$. We set $N_z = 50$ and verify that increasing N_z does not change results materially.

2. Start with initial values for wages and the bond interest rate $\{w_i\}_{i=1}^N$, r^b.

3. Solve value function $\hat{v}_i$ as the fixed point of equation (B.10) using value function iteration. The integral over z' is evaluated using the trapezoidal rule over the predetermined grid to calculate the expected continuation value. The solution also gives policy functions for consumption and saving.[22]

4. Apply the policy functions for saving and the expected return on net worth to solve the wealth density function ϕ_i as the fixed point of equation (B.3). In doing so we use an iterative procedure and the trapezoidal rule over the predetermined grid to evaluate the integral over z.

   One challenge for this integration is that policy functions are discontinuous at the threshold productivity levels at which parent firms switch from inactive to active and affiliates switch from zero to positive leverage. To overcome this challenge, we exploit the fact that when the firm distribution can be represented by continuous density functions, the set of firms at these threshold productivity levels has a zero mass. For any productivity level $\hat{z}$ at which policy functions are discontinuous, we

[21] This estimate is not far off from using equation (C.2) as the structural equation directly, in which case we would obtain $\hat{\eta}_z = 0.184/\theta = 0.0368$.

[22] In the baseline model with log utility, the policy function for saving can be derived without solving the value function so this step can be skipped if the value function is not of interest. In settings with general CRRA utility function or firm-level adjustment costs, it is necessary to use the value function iteration for solving the policy functions.

add $\hat{z}-\varepsilon$ and $\hat{z}+\varepsilon$ to the set of integration nodes with $\varepsilon = 1e-12$. Outside regions $(\hat{z}-\varepsilon, \hat{z}+\varepsilon)$, policy functions are continuous, so the standard trapezoidal rule performs well; within these intervals, policy functions are discontinuous but with sufficiently small ε, the measures of these intervals are arbitrarily close to zero. The discontinuity in policy functions therefore does not lead to discontinuity of aggregate variables in prices or parameters, ensuring numerical stability of our algorithm. This demonstrates the appeal of working on the density functions of the firm distribution.

5. Compute the aggregate labor supply, bond supply, bond demand from parent firms, and bond demand from affiliates, following the characterizations in Section B.1. For integration which involves policy functions that are discontinuous in the productivity level, apply the strategy described in Step 4. Check whether the labor markets and the global bond market clear. If both markets clear, the prices of the stationary equilibrium have been solved. Otherwise, update wages and bond interest rate and go back to Step 2.

Calibration for the steady state. The calibration uses a nested procedure. In the inner loop, we solve each country's fundamental productivity level $\bar{z}_i$, financial constraint parameters of the parent firm λ_i and of the affiliate μ_h, and the bilateral FDI return wedges $\bar{\eta}_{ih}, \forall h \neq i$, such that at steady state the model matches the data in year 2001 on the following moments: (1) GDP per efficient labor in each country; (2) credit-GDP ratio in each country; (3) the share of affiliates' assets financed by parents in each host country; (4) bilateral FDI stocks as shares of each host country's capital stock. In the outer loop, we set η_z according to the procedure that is described in Section C.2.1.

Transition path. We use a shooting method to solve the transition path.

1. Choose the same predetermined grid for productivity z as the one used in solving the steady state.

2. Choose a large finite horizon T.

3. Start with initial values for the sequence of wages and bond interest rate $\{\{w_{i,t}\}_{i=1}^{N}, r_t^b\}_{t=0}^{T}$.

4. Solve the sequence of value functions $\{\hat{v}_{i,t}\}_{t=0}^{T}$ following equation (B.10) backwardly from period T, assuming $\hat{v}_{i,T+1}$ equals the value function at the future steady state. The solution also generates a sequence of policy functions for consumption and saving.

5. Solve the sequence of wealth density functions $\{\phi_{i,t}\}_{t=0}^{T}$ following equation (B.3), starting from the initial wealth density function $\phi_{i,0}$. For integration involving policy functions that are discontinuous in the productivity level, we apply the same strategy that is described for solving the stationary equilibrium.

6. Compute aggregate quantities based on the policy functions and the wealth density functions. Check whether the labor markets and the global bond market clear in each period. If not, update the sequence of wages and bond interest rate and go back to Step 4.

7. Check whether the implied wealth density functions are stationary at period T and $\hat{v}_{i,T}$ is closed to the value function at the future steady state. If not, increase T and go back to Step 3.

Calibration for the transition path. We solve the sequence of each country's fundamental productivity level $\bar{z}_i$, leverage constraints for parents(λ_i) and affiliates (μ_h), domestic and FDI investment return wedges $\bar{\eta}_{ih}$, to match data on the following moments from year 2001 to 2012: (1) GDP per efficient labor in each country; (2) credit-GDP ratio in each country; (3) the share of affiliates' assets financed by parents in each host country; (4) each country's capital stock and the bilateral FDI stocks as shares of each host country's capital stock. We assume that the parameters remain at their 2012 values for the years after, so the transition path converges to a steady state.[23]

C.2.3 Additional Results from the Calibration

Table C.1 reports some country characteristics and the corresponding parameters. Columns 1, 3, 5 are the data—the average values of the credit over GDP ratio, the ratio between total finance and parent finance, and the share of inward FDI stock in a host's total capital stock, respectively. Columns 2, 4, 6 report the average values of the parameters that are pined down by these data moments.

C.3 The GE Effect and Validation from Country-Specific Experiments

The decomposition reported in Section 5.2 of the main text focuses on the world FDI in response to changes in a certain set of fundamentals of all countries at the same time. In this section, we conduct country-specific experiments, which serve two purposes. First, by comparing the results from country-specific experiment and the result in Section 5.2 we can shed light on the importance of the general equilibrium effects. Second, by comparing the results from these country-specific experiments to the dynamic patterns seen in the data, we further validate the model.

FDI growth, 2001-2007. We first examine the extent to which the easing access to credit in the lead-up to the financial crisis accounts for the increase in FDI during this period for individual countries. Each bar in Figure C.3 represents the experiment for one country; the height of the bar corresponds to the actual *cumulative* net outward FDI *flows* from 2002 to 2007—or equivalently, the level increase in outward FDI stock from 2001 to 2007. Consistent with the increase in aggregate FDI shown in Figure 4, most countries see increasing outward FDI stock during the sample period.

We decompose the cumulative outflow from each country into four components. For experiments on country i, we change only the targeted parameter of country i, keeping all other parameters for country i and the rest of countries at the calibrated values. We add back time subscripts to variables when necessary to highlight their time dependence.

In the first set of experiments, we set $\lambda_{i,t}$ to $\lambda_{i,2001}$ for country i and solve for the counterfactual transitional path, one country at a time. The pink solid bars in Figure C.3 demonstrate the strength of this force. More precisely, the height of the pin solid bar indicates by how much the outward FDI would have been lower, had $\lambda_{i,t}$ stayed at $\lambda_{i,2001}$. A positive value indicates the change in $\lambda_{i,t}$ between 2001 and 2007 increases the growth in outward FDI. Although the importance of this channel differs, for most countries, the contribution is positive.

[23]For the baseline model with log utility, given the distribution of firms at initial period, the calibration (as well as the determination of equilibrium wages and bond interest rate) for a given period can be done independently of future periods, since the policy functions, as characterized in Lemma 2, do not depend on the continuation value.

Table C.1: Country Characteristics and Calibrated Parameters

ISO	(1) Credit/GDP	(2) $\bar{\lambda}_i$	(3) Host Leverage	(4) $\bar{\mu}_h$	(5) FDI Share	(6) $\bar{\eta}_{ih}$
ARG	0.13	0.07	1.58	0.73	0.06	0.51
AUS	1.12	0.77	2.03	1.44	0.06	0.53
AUT	0.93	0.23	1.67	1.16	0.04	0.50
BEL	0.61	0.11	1.53	0.92	0.26	0.73
BRA	0.41	0.10	1.77	0.93	0.04	0.46
CAN	1.84	1.24	2.42	1.81	0.09	0.58
CHE	1.52	0.67	1.43	0.64	0.15	0.61
CHL	0.87	0.78	1.71	0.84	0.10	0.38
CHN	1.17	0.61	1.95	1.00	0.01	0.29
CZE	0.37	0.07	1.56	0.96	0.05	0.47
DEU	0.99	0.31	1.99	1.42	0.03	0.74
DNK	1.69	0.49	1.62	1.02	0.05	0.55
ESP	1.43	0.30	1.67	1.19	0.04	0.55
FIN	0.74	0.17	1.86	1.40	0.04	0.60
FRA	0.86	0.16	1.94	1.47	0.06	0.76
GBR	1.58	0.35	1.73	1.05	0.12	0.72
HUN	0.48	0.12	1.22	0.32	0.11	0.43
IDN	0.25	0.03	1.35	0.40	0.06	0.32
IND	0.43	0.17	1.74	0.76	0.01	0.18
IRL	1.22	1.00	1.48	0.68	0.55	0.61
ITA	0.78	0.15	1.87	1.56	0.02	0.59
JPN	1.79	0.79	2.36	1.88	0.01	0.49
KOR	1.27	0.45	1.89	1.21	0.02	0.48
MEX	0.20	0.07	1.78	0.90	0.05	0.44
MYS	1.11	0.92	1.55	0.65	0.07	0.37
NLD	1.14	0.49	1.47	0.71	0.44	0.81
NOR	1.12	0.70	1.76	0.99	0.07	0.48
NZL	1.27	1.14	1.72	0.90	0.13	0.46
POL	0.35	0.23	1.71	0.78	0.06	0.39
PRT	1.38	0.29	1.69	1.51	0.04	0.47
RUS	0.33	0.29	1.25	0.29	0.01	0.31
SGP	1.00	0.31	1.62	1.01	0.22	0.51
SWE	1.09	0.30	1.82	1.22	0.07	0.60
TUR	0.29	0.12	1.73	0.82	0.02	0.28
USA	1.84	1.41	2.07	1.55	0.03	0.95
VEN	0.17	0.05	1.72	0.96	0.04	0.40
Mean	0.94	0.43	1.73	1.03	0.09	0.52
Std	0.51	0.37	0.25	0.38	0.11	0.16

$\bar{\lambda}_i$ and $\bar{\mu}_h$ are averaged over time for each country. $\bar{\eta}_{ih}$ is the inward FDI return wedge, averaged across home country i and over time t. See the text in Section C.2.3 for additional explanation.

To isolate the effect of a change in affiliate financing for a single sending country, we abuse the notation by denoting μ_{ih} the affiliate financing constraint of firms from home country i in host country h, allowing μ_{ih} to be different across home country i. We then set $\mu_{ih,t}$ for affiliates from home country i to its 2001 value, for all $h \neq i$, while keeping $\mu_{i'h,t}, i' \neq i$ at the benchmark values. This exercise captures the impact on FDI through the 'pull' force of increasing credit availability in a host country. The green striped bars in Figure C.3 show that, by making it easier for parent companies to access external finance in the host country, financial market conditions elsewhere could have a quantitatively significant impact on the investment decisions of MNEs.

In the last set of experiments, we explore the influence of domestic productivity growth on outward

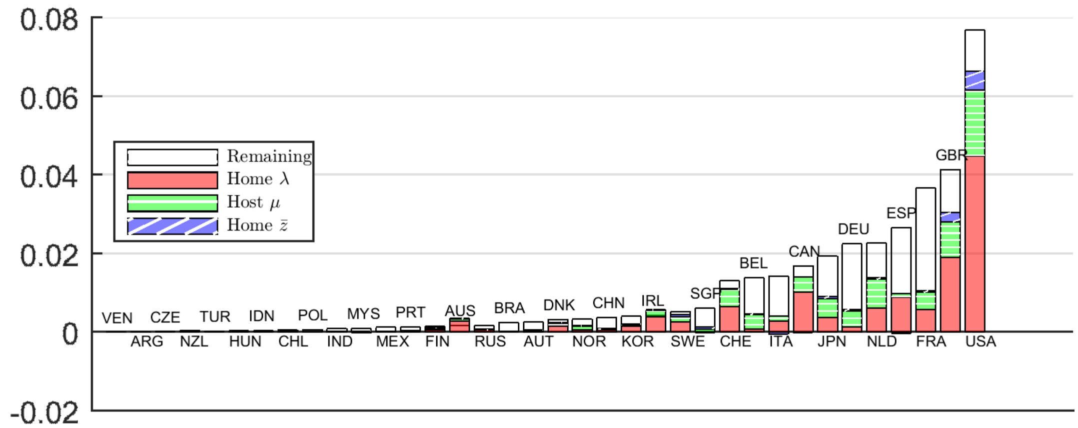

Figure C.3: Cumulative Outward FDI Flows: 2002-2007

Note: Decomposition by country, 2002-2007. The height of the bar indicates the total cumulative outward FDI flow for each country. The colored bars indicate fractions accounted for by individual channels..

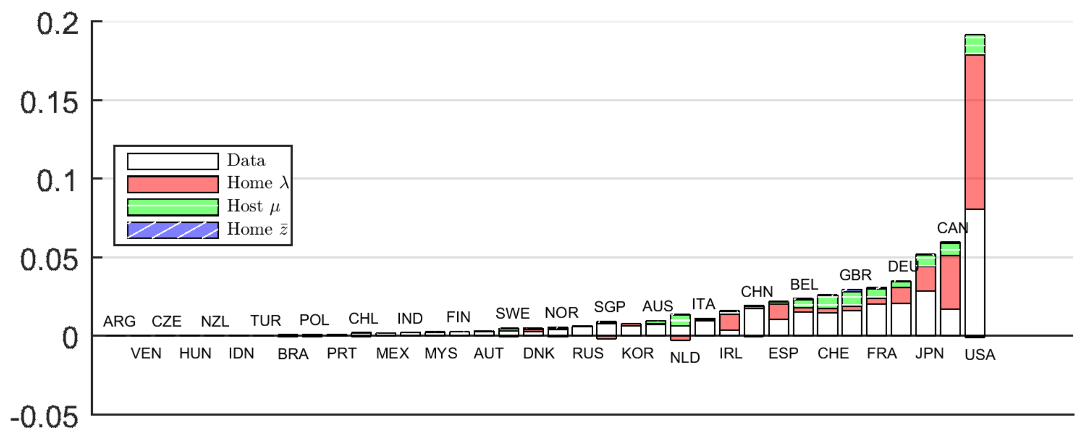

Figure C.4: Cumulative Outward FDI Flows: 2008-2012

Note: Decomposition by country, 2008-2012. The white bars indicate the factual cumulative outward FDI flows for each country; the colored bars indicate the predicted *additional* outward FDI due to changes in different set of parameters.

FDI. The blue shaded bars in the figure indicate the importance of this channel, which differs significantly across countries. For the U.S., domestic productivity growth plays a somewhat important role; this is not the case in many European countries, such as Spain and France. This heterogeneity primarily reflects the difference in TFP growth rate across countries.

The white bars in Figure C.3 are the remaining cumulative FDI outflow during this episode after deducting the above three channels. This term encompasses changes in the investment and FDI return wedges—which are not formally modeled and could be driven by policy, technology, or misspecifications in the model—as well as the interaction among countries.

Growth slowdown, 2008-2012. We now investigate the role of the credit crunch during and immediately after the financial crisis on the slowdown in FDI from each country. As before, we perform three sets of experiments. Instead of keeping $\lambda_{i,t}$, $\mu_{ih,t}$, and $\bar{z}_{i,t}$ constant at their 2001 values, however, we feed in the calibrated values until 2007 and fix them afterwards. The question we ask in these experiments is, for example, how much *more* outward FDI we would see, had $\lambda_{i,t}$ stayed at the 2007 peak value.

Figure C.4 presents the results. The white bars show the factual cumulative FDI outflow during 2008-2012. The red bars show the additional outward FDI from country i, if $\lambda_{i,t}$ stays at the value of 2007 for

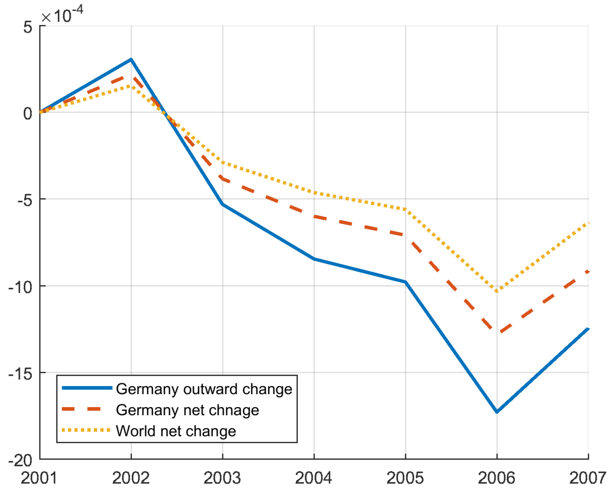

Figure C.5: The Importance of General Equilibrium Effects: the case of Germany over 2002-2007

subsequent periods. For the U.S., the impact is particularly stark. In countries whose financial market was less interrupted by the crisis, such as China, this counterfactual barely makes any difference.

The green stripped bars show that, disruptions in the financial market of destination countries reduce the incentive for foreign MNEs to invest. As the biggest sending country of FDI, the U.S. is the most affected, but this channel is also important for Netherlands, Switzerland, and the U.K., which invest heavily in other countries that were severely affected by the European debt crisis. Finally, the blue shaded bars show that the role of domestic productivity is negligible in most countries.

The importance of general equilibrium effects. The above counterfactual exercises show that the changes in financial market conditions can have significant impacts on the dynamics of outward FDI from individual countries. If we had simply aggregate the impacts of individual countries based on country-specific experiments, findings in Figure C.3 and C.4 suggest that financial factors can explain more than half of the cumulative FDI outflow during 2002-2007; in addition, had the access to credit remained at the peak level of 2007, the cumulative FDI outflow during 2008-2012 could almost double.

The sum of country-specific effects overestimates by a factor of two the actual effects, reported in Section 5.2 of the main text. To illustrate the source of the difference through an example, Figure C.5 plots several variables from an experiment that fixes $\lambda_{i,t}$ at $\lambda_{i,2001}$ for Germany throughout. The solid line indicates that, without the improvement in the financial market, the cumulative FDI outflow from Germany would decrease. Because the credit boom in Germany was only moderate, by the end of 2007, the total decrease is around -0.0013 (U.S. 2001 GDP normalized to 1), which accounts for around 8% of the cumulative FDI outflow from Germany during this period. The level 0.0013 corresponds to the height of the solid pink bar for Germany in Figure C.3.

This change overstates the influence of German financial shock on the *world* FDI for two reasons. The first is due to a domestic general equilibrium effect: in the absence of the credit boom, the wage in Germany decreases. This attracts more inward FDI, so the aggregate FDI in the world decreases by less. The second is due to a third-country effect. As German firms decrease their overseas investment, the labor market in these destinations become less competitive, drawing more investment from third countries. The dashed line in Figure C.5 plots the net change in the sum of inward and outward FDI in Germany; the dotted line plots the net change in the sum of FDI across all country pairs. The difference

Table C.2: Diff-in-Diff Estimate of Home Financial Market on Outward FDI

	Data		Model	
	(1)	(2)	(3)	(4)
$\Delta\log(\text{Credit/GDP})$	0.698***		0.375***	
	(0.155)		(0.092)	
$\Delta\log(\lambda_i)$		0.294**		0.282***
		(0.132)		(0.046)
Year FE	yes	yes	-	-
Observations	364	364	396	396
R^2	0.205	0.194	0.375	0.577

Note: This table reports the effect of home financial market conditions on outward FDI using panel data. The dependent variable is the yearly change in log outward FDI stock. The independent variables are the yearly change in credit over GDP ratio or λ_i. The first two columns are estimated based on the actual data, with year fixed effects controlled for. The last two columns are estimated based on model counterfactuals, where both dependent and independent variables are the log difference between the benchmark and counterfactual variables and then first-differenced.
Standard errors (clustered by country) in parenthesis.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

between the solid and dashed lines shows the strength of the first force; the difference between the dotted and dashed lines is due to the second force. In the case of the German λ_i shock, both forces play a similar role; together, they reduce the inferred decline in the aggregate FDI by half.

The difference between the dotted line and the solid line demonstrates that to understanding the dynamic effects of financial market conditions on world FDI, it is crucial to incorporate the general equilibrium mechanisms.

Validating dynamic implications using diff-in-diff estimates. The results from the above counterfactual exercises naturally serve as the basis for a validation test. The idea is as follows: we can estimate the effect of credit market conditions on outward FDI through diff-in-diff (subject to the caveat in identification). The above counterfactual exercises give us the 'treated version' of each country, which we can compare to the 'control' group—the actual data—for reduced-form estimation. Comparing the diff-in-diff estimates and the estimates based on the counterfactuals provides a test for whether the model speaks to the dynamic patterns in the data. It is a strong test of the model because it is based on outcomes from counterfactuals, rather than objects we directly or indirectly target in calibration.[24]

The first two columns of Table C.2 report this estimate based on the actual data. The dependent variable is log outward FDI stock. The independent variables are the credit over GDP ratio and the calibrated λ_i. Variables are first-differenced and year fixed effects are controlled for. The estimates suggest that home country financial market conditions have a positive and statistically significant correlation with outward FDI. The estimated coefficient is larger when financial market conditions are measured using credit over GDP than when they are measured using the calibrated λ_i.

The third and fourth columns report the corresponding estimate based on model counterfactuals. Specifically, for each country i, we first calculate the differences in its outward FDI and measures of financial market conditions between the baseline economy and the *counterfactual* economy in which λ_i is altered in ways described above. Effectively, this difference nets out the year fixed effects; we then first-difference these variables and estimate an OLS specification, to obtain a diff-in-diff estimate for the

[24]Note that because our calibration fits the data perfectly, if we had use the equilibrium outcomes in the model to perform the regression, by design we will obtain exactly the same estimates.

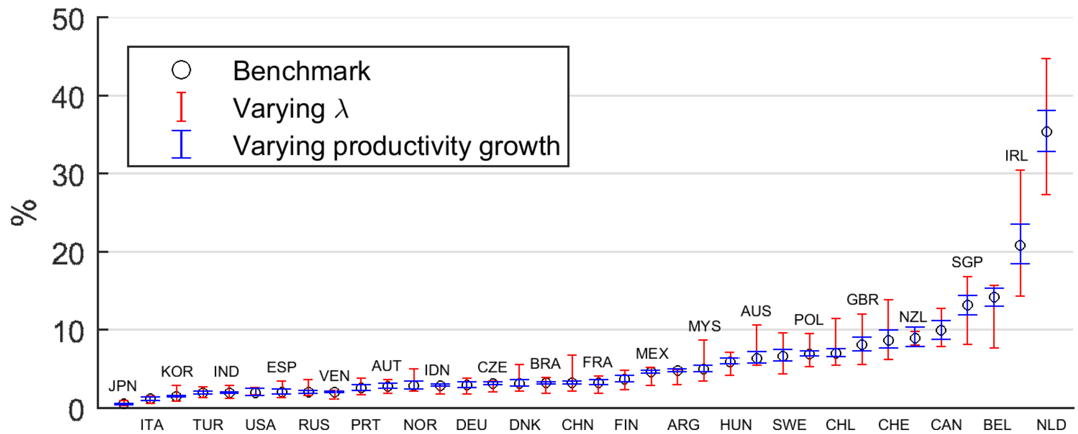

Figure C.6: Average Dynamic Wage Gains under Counterfactual Primitives

Note: Plotted are dynamic wage gains defined as the average change in log wages for a host country during 2001-2012, moving the country from a counter-factual economy with its inward FDI shutdown to the calibrated economy. 'Benchmark' corresponds to the the baseline calibration. 'Varying λ' corresponds to the range of average dynamic wage gain when host country h's financial development parameter λ_h varies by 2 s.d. (0.76) in both directions. 'Varying productivity growth' corresponds to the range when the host country's annual productivity growth rate varies by 2 s.d. (5%) in both directions.

effect of a change in home country λ_i on outward FDI.

We find that the elasticities for credit over GDP and (especially) λ_i are in the same order of magnitude as those reported in columns 1 and 2, so our model is able to replicate the reduced-form regressions. Note that in the model, financial market conditions affect both contemporaneous and future FDI. Thus, the diff-in-diff estimates using the simulated data captures both the contemporaneous effect and the cumulated effects from past financial market conditions. That these estimates are close to their empirical counterparts suggests that our model captures well the dynamic relationship between financial market conditions and FDI in the data.

C.4 The Role of Host Fundamentals in Shaping Dynamic Wage Gains

Section5.3 of the text evaluates the welfare gains from openness to FDI and explores the roles of countries' fundamentals in determining the dynamic gains using the example of Hungary. We conduct similar exercises for each sample country and summarize the results in Figure C.6. The circles denote the average dynamic wage gains in the baseline, corresponding to Column 4 of Table 6. The intervals denote the range of dynamic wage gains as we vary host country h's financial development parameter λ_h from $\lambda_h - 0.76$ to $\lambda_h + 0.76$, and the growth rate of its fundamental productivity ($\bar{z}_h$) from the baseline to plus and minus 5%. What stands out is the wide range of possible outcomes, especially as we vary λ_h. For Malaysia and Austria, for example, the upper end of the range is about twice as much as the baseline values; for Mexico, the baseline value is close to the upper end, but much higher than the bottom of the range. Such heterogeneity, again, highlights the crucial role of country fundamentals in shaping the dynamic welfare effects.

References

Antras, Pol, Mihir A Desai, and C Fritz Foley, "Multinational firms, FDI flows, and imperfect capital markets," *The Quarterly Journal of Economics*, 2009, *124* (3), 1171–1219.

Arkolakis, Costas, Andrés Rodríguez-Clare, and Jiun-Hua Su, "A multivariate distribution with pareto tails," *Mimeo*, 2017.

Bloom, Nicholas, Christos Genakos, Raffaella Sadun, and John Van Reenen, "Management practices across firms and countries," *The Academy of Management Perspectives*, 2012, *26* (1), 12–33.

Bollard, Albert, Peter J Klenow, and Huiyu Li, "Entry Costs Rise with Development," *Mimeo*, 2016.

Cravino, Javier and Andrei A. Levchenko, "Multinational firms and international business cycle transmission," *The Quarterly Journal of Economics*, 2017, *132* (2), 921–962.

Fan, Jingting, "Talent, Geography, and Offshore R&D," *Manuscript*, 2017.

Guvenen, Fatih, Jr. Mataloni Raymond J, Dylan G Rassier, and Kim J Ruhl, "Offshore profit shifting and domestic productivity measurement," NBER Working Paper 23324 2019.

Nocke, Volker and Stephen Yeaple, "An assignment theory of foreign direct investment," *The Review of Economic Studies*, 2008, *75* (2), 529–557.

Ramondo, Natalia, Andrés Rodríguez-Clare, and Felix Tintelnot, "Multinational production: Data and stylized facts," *The American Economic Review*, 2015, *105* (5), 530–536.

Torslov, Thomas R, Ludvig S Wier, and Gabriel Zucman, "The Missing Profits of Nations," Working Paper 24701, National Bureau of Economic Research June 2018.