Online Appendix for "Multinational Production with Non-neutral Technologies": Not for Publication

Chang Sun

July, 2019

OA.1 Data

In this data appendix, I provide additional details on the data used in the paper as well as on how I infer missing values.

OA.1.1 Aggregate Variables

Aggregate variables are used in the firm-level regressions as well as in the quantitative implementation.

Total Nonfinancial Output

I obtain total nonfinancial output from World KLEMS and OECD STAN for each country. If a country is covered by both databases, I use the one with better coverage in terms of years. For country-years with missing data, I extrapolate using a log-linear relationship between total nonfinancial output and GDP, following Ramondo et al. (2015). In particular, I run the following regression for three different periods, including the baseline period 1996–2001 and the counterfactual period 2006–2011:

$$\log Y_{lt} = \alpha_0 + \alpha_1 \log GDP_{lt} + u_{lt}. \tag{1}$$

As is shown in Table OA.1, the coefficients before GDP are very close to 1 in all three periods. The constant, which reflects the log of output-GDP ratio when $\alpha_1 = 1$, appears to have declined over time. This suggests that intermediate inputs have become more important in production. Among the 37 sample countries, nine of them do not have any data on total output in 1996–2001 and 2006–2011, so the average total output is calucated based on extrapolated values only. These countries are: Argentina, Bulgaria, Brazil, the Dominica Republic, India, Romania, Turkey, Uruguay and Venezuela.

Table OA.1: Estimate extrapolation equations for different periods

Year Range	96-01	02-05	06-11
log(GDP)	0.984	0.988	1.003
	(0.0113)	(0.0136)	(0.0171)
Cons.	0.993	0.889	0.546
	(0.305)	(0.372)	(0.463)
N	212	144	96
R^2	0.983	0.982	0.982

Dependent variable is log of total nonfinancial output. Pooled regression for all country-years in the sample. Robust standard errors in parentheses.

Endowment

Countries' endowment in capital and labor comes from the Penn World Table 8.0. I use capital stock at constant 2005 national prices (variable "rkna", in mil. 2005US$) to measure capital endowment. For labor endowment, I multiply the number of persons engaged (variable "emp", in mils) by the average human capital (variable "hc") to obtain efficiency units of labor. The Penn World Table 8.0 uses the Barro-Lee dataset and constructs the human capital combining the average years of schooling in each country and assumed return to schooling (see Robert Inklaar and Timmer (2013) for the methodology to develop these variables).

Labor Shares

Karabarbounis and Neiman (2014) provide two measures of labor shares: total labor shares and corporate labor shares. The corporate labor share is the labor share within the corporate sector, and is not influenced by the methods of imputing the labor shares for the non-corporate sectors. It is also more consistent with my micro evidence on multinational firms' technologies. Therefore, I use the corporate labor share wherever possible. For the baseline period in the counterfactuals, all 37 sample countries have data on corporate labor shares. In the later period (2006–2011), seven countries lack information on corporate labor shares; thus I use the total labor shares when needed. Factor prices (wages for one efficiency unit of labor and rental rates) are backed out using GDP, labor shares and endowment.

Bilateral Trade Data

Bilateral trade data ($\sum_i X_{iln}$) are available from the BACI database (Guillaume and Zignago (2010)). I use the data to calculate the bilateral trade shares as well as total absorption $X_n = \sum_{i,l} X_{iln}$, where the domestic trade volumes are backed out using total output and total export $\sum_i X_{ill} \equiv Y_l - \sum_{i,n\neq l} X_{iln}$.

Bilateral MP Data 1996-2001

Ramondo et al. (2015) provide bilateral MP sales ($\sum_n X_{iln}$ as in the model) for all country pairs in my sample for the baseline period (1996–2001). This allows me to calibrate the full MP cost matrix γ_{il} in the baseline calibration.

Bilateral MP Data 2006-2011

For the first counterfactual experiment, I need data on multinational sales for the later period, 2006–2011. The data is constructed using OECD and Eurostat Foreign Affiliates Statistics (FATS) databases and averaged over the period of six years. Some countries may have missing values in certain years. Here I discuss the extrapolation procedures.

For 22 countries in my first counterfactual experiment (excluding China), OECD and Eurostat contain at least two nonmissing total inward multinational sales between 2006 and 2011. Among the total $22 \times 6 = 132$ observations between 2006 and 2011, only 14 have missing values. I extrapolate these missing values using the following two methods.

1. Predict missing total inward MP sales by the following regression:

$$\log X_{lt} = \beta \log Q_{lt} + \delta_l + \delta_t + \epsilon_{lt},$$

   where X_{lt} is country l's inward MP sales in year t, Q_{lt} is the total inward FDI stock, and δ_l and δ_t are country and year fixed effects, respectively. The sample of the regression includes all countries covered by the Eurostat or OECD databases with at least three observations between 2006 and 2012. The estimated coefficient for β is 0.1, and the regression has an R-squared of 0.998. This step extrapolates 10 out of the 14 missing values for inward MP sales.

2. Predict missing total inward MP sales using a linear time trend. In particular, I run the following regression

$$\log X_{lt} = \delta_l + \delta_l \times t + \delta_t + \epsilon_{lt},$$

   where $\delta_l \times t$ is country-specific time trends while other terms are defined as in step 1. This extrapolates the remaining four missing values.

For China, I use the industrial enterprises database to compute the total output by foreign-owned manufacturing firms (excluding firms owned by Hong Kong, Macau and Taiwan investors) in each year. I then scale the foreign-owned manufacturing firms' total output by the share of FDI in manufacturing to obtain an estimate for the total output by all foreign-owned firms in China.

For 21 out of the 23 countries in our baseline counterfactuals, Eurostat-OECD contains information on bilateral MP sales for at least three years within this six year period. I take the average bilateral MP shares (MP sales divided by total output of the host country) and use them for calibrating the new MP costs.

OA.1.2 Construction of Asset Deflators

Since I only use the 2012 firm-level data, I cannot perform a perpetual inventory method to calculate the real stock of capital. Here I provide a way to construct an asset deflator under assumptions about the growth rates of investment prices, the growth rate of firm-level capital stocks and the number of years that the firms have been accumulating capital. Consider the perpetual inventory method for a typical firm in country l:

$$\tilde{K}_{lt} = I_{lt} + \frac{P_{lt}^I}{P_{lt-1}^I} (1 - \delta_l) \tilde{K}_{lt-1},$$

where I_{lt} is the **value** of investment in period t at the price of P_{lt}^I and $\tilde{K}_{lt}$ is the **value** of capital stock at the price of P_{lt}^I, at the end of period t. δ_l is the country specific discount rate. Iterating backwards, we have

$$\begin{aligned} \tilde{K}_{lt} &= I_{lt} + \frac{P_{lt}^I}{P_{lt-1}^I} (1 - \delta_l) \tilde{K}_{lt-1} \\ &= \sum_{j=0}^{\infty} (1 - \delta_l)^j \frac{P_{lt}^I}{P_{lt-j}^I} I_{lt-j}. \end{aligned}$$

The real stock of capital equals the current value of capital stock deflated by the current investment price:

$$K_{lt} \equiv \frac{\tilde{K}_{lt}}{P_{lt}^I} = \sum_{j=0}^{\infty} (1 - \delta_l)^j \frac{1}{P_{lt-j}^I} I_{lt-j}.$$

In practice, however, the book value of past investment are not adjusted as investment prices change over time. I approximate the book value of total assets as follows

$$\tilde{K}_{lt}^{acct} \equiv I_{lt} + (1 - \delta_l) \tilde{K}_{lt-1}^{acct} = \sum_{j=0}^{\infty} (1 - \delta_l)^j I_{lt-j}.$$

Note that in Crozet and Trionfetti (2013), the authors deflate the accounting value of capital stock using the price of investment and obtain

$$K_{lt}^{CT} \equiv \tilde{K}_{lt}^{acct} / P_{lt}^I = \sum_{j=0}^{\infty} (1 - \delta)^j \frac{1}{P_{lt}^I} I_{lt-j},$$

which tends to underestimate the real capital stock if there is constant inflation in investment prices. To properly adjust the real capital stock, I assume that all countries are on balanced growth paths and *the real capital stock grows at a constant rate* g_l. This implies

$$
\begin{aligned}
K_{lt} &= (1+g_l)\,K_{lt-1} = I_{lt}/P_{lt}^I + (1-\delta_l)\,K_{lt-1} \\
&\Rightarrow \; I_{lt}/P_{lt}^I = (g_l+\delta_l)\,K_{lt-1}.
\end{aligned}
$$

Thus real investment grows at the same speed as capital stock. Also assume the investment prices grow at constant rates π_l. I can rewrite the real capital stock as

$$
\begin{aligned}
K_{lt} &= \sum_{j=0}^{\infty} (1-\delta_l)^j \frac{1}{P_{lt-j}^I} I_{lt-j} \\
&= \frac{I_{lt}}{P_{lt}^I} \sum_{j=0}^{\infty} \left(\frac{1-\delta_l}{1+g_l}\right)^j \\
&= \frac{I_{lt}}{P_{lt}^I} \frac{1+g_l}{\delta_l+g_l},
\end{aligned}
$$

and the book value of capital stock

$$
\begin{aligned}
\tilde{K}_{lt}^{acct} &= \sum_{j=0}^{\infty} \left(\frac{1-\delta_l}{1+g_l}\right)^j \frac{I_{lt}}{P_{lt}^I} P_{lt-j}^I \\
&= \frac{I_{lt}}{P_{lt}^I} \sum_{j=0}^{\infty} P_{lt}^I \left(\frac{1-\delta_l}{(1+\pi_l)(1+g_l)}\right)^j \\
&= I_{lt} \frac{(1+\pi_l)(1+g_l)}{(1+\pi_l)(1+g_l)-(1-\delta_l)}.
\end{aligned}
$$

Thus the proper deflator is

$$
\frac{\tilde{K}_{lt}^{acct}}{K_{lt}} = P_{lt}^I \frac{(1+\pi_l)(\delta_l+g_l)}{\pi_l+g_l+\pi_l g_l+\delta_l}.
$$

If the life span of firms is finite, say T, then the deflator should be

$$\begin{aligned}\frac{\tilde{K}^{acct}_{lt}}{K_{lt}} &= P^I_{lt}\frac{\sum_{j=0}^{T-1}\left(\frac{1-\delta_l}{(1+\pi_l)(1+g_l)}\right)^j}{\sum_{j=0}^{T-1}\left(\frac{1-\delta_l}{1+g_l}\right)^j} \\ &= P^I_{lt}\frac{(1+\pi_l)(\delta_l+g_l)}{\pi_l+g_l+\pi_l g_l+\delta_l}\frac{1-\left(\frac{1-\delta_l}{(1+\pi_l)(1+g_l)}\right)^T}{1-\left(\frac{1-\delta_l}{1+g_l}\right)^T}.\end{aligned}$$

In practice, I calculate g_l and π_l using the average growth rates of aggregate capital stock and investment prices for the period 1990–2011. I set firm age T in all countries at a particular value. The choices of g_l, π_l and T do not matter in most of the reduced-form regressions, since the deflators constructed here are country-specific, and they are absorbed by host-country-industry fixed effects. However, when it comes to the estimation of the intensive elasticity, the identification comes from the variation in aggregate factor prices and firms' capital-labor ratios across host countries. Therefore, the choices of asset deflator will affect the estimates for the intensive elasticity. For the baseline estimate, I assume firm age to be 10 years in all countries. I experiment with different values of T to construct the asset deflator and the estimates are similar.

OA.1.3 Firm-level Data and Robustness to Data Cleaning and Coverage

The firm-level data are a cross-section for 2012, downloaded from the Orbis online interface in July 2014. To study firms' capital intensity, I start with firms' unconsolidated accounts with nonmissing key variables for the regressions (host country, industry, total assets, wage bills and number of employees).

I then use four steps to obtain the final sample (Table OA.2 shows the number of observations dropped in each step)

1. Drop firms in the financial sector (NACE Rev 2 Sectors 64, 65 and 66). This step drops 2.5% of the original sample.

2. Drop firms that are located or headquartered in tax havens (Gravelle, 2009). The major tax haven countries in my original sample are Switzerland, Luxembourg, Ireland and Cyprus. This step drops 0.3% of the original sample.

3. Drop observations with abnormal values of capital, labor and capital intensities. In particular, I first drop observations if its total asset, employment or wage bill are

zero or negative. Next, I calculate the capital-to-employment and capital-to-wage-bill ratios, and drop observations whose ratios are above 200 times or below 1/200 of the country-industry median, following the practice in Crozet and Trionfetti (2013). I also drop firms whose cost per worker is above one million USD or above the 98th percentile or below the 2nd percentile of the wage distribution within a country-year. This step drops another 0.3% of the sample. Though these thresholds are somewhat arbitrary, I show in columns 1 and 2 of Table OA.3 that the technology and size effects are robust to keeping these outliers in the sample.

4. Drop countries and industries with too few observations. In particular, I drop home and host countries with fewer than 5000 firms. I drop industries if they span fewer than 5 countries with at least 50 observations in each country. This step again follows the practice in Crozet and Trionfetti (2013), though the threshold number of firms for countries to be included (5000) is higher than theirs (300). This is because I want to make sure that the technologies origin effects are not driven by home countries with few firms. This step drops about 4% of the sample.

   In columns 3 and 4 of Table OA.3, I use less stringent criteria to select an alternative sample. I drop home and host countries with fewer than 2000 firms and drop industries if they span fewer than 3 countries with at least 50 observations in each country. This criteria adds back 1.8% of the original sample back. The technology origin and size effects are both robust when estimated using the alternative sample.

Table OA.2: Data cleaning procedures

Steps	# of obs	% of original sample
Firms with non-missing key variables	2822736	100.0
Step 1: drop financial firms	70164	2.5
Step 2: drop MNE in tax havens	9590	0.3
Step 3: drop firms with abnormal capital intensities	9328	0.3
Step 4: drop countries or industries with too few firms	112099	4.0
(Step 4': a version of Step 4 with less stringent requirements)	(62069)	2.2
Final sample	2621555	92.9

0. Key variables refer to: total assets, number of employees, cost of employees, host country, industry and home country K/hL.

1. Financial firms refer to firms in the financial sector (NACE Rev2 Sectors 64, 65 and 66).

2. Drop affiliates in and from tax havens (Gravelle, 2009). The major tax haven countries in my original sample are Switzerland, Luxembourg, Ireland and Cyprus.

3. An observation is identified with abnormal K/L or K/wl if any of the three variables is non-positive, or the ratio is larger than 200 times of the country-industry median, or smaller than 1/200 of the country-industry median. I also drop firms whose cost per worker is above one million USD or above the 98th percentile or below the 2nd percentile of the wage distribution within a country-year.

4. Host and home countries are dropped if fewer than 5000 firms are in the dataset. Industries are dropped if they span fewer than 5 countries with at least 50 observations in each country.

4.' In the less stringent version of Step 4 (Step 4'), I drop host and home countries with fewer than 2000 firms. Industries are dropped if they span fewer than 3 countries with at least 50 observations in each country.

Table OA.3: Size Technology Origin Effects: Including Outliers

	Dependent Var: log(total assets/wage bill)			
	Including Outliers in K/wL		Further Including Cty/Ind with fewer firms	
	All Firms (1)	MNE (2)	All Firms (3)	MNE (4)
$\log(K_i/L_i)$	0.189***	0.266*	0.178**	0.261*
	(0.0690)	(0.145)	(0.0664)	(0.142)
log(revenue)	0.0441**	0.0396***	0.0443**	0.0415***
	(0.0215)	(0.0100)	(0.0212)	(0.00940)
debt-to-equity ratio	0.00405***	0.00455***	0.00402***	0.00416***
	(0.00125)	(0.000752)	(0.00125)	(0.000711)
# of firms	2,011,000	45,000	2,052,000	52,000
# of host countries	24	22	24	23
# of host * industry	7,333	3,740	8,802	4,402
# of home countries	40	39	40	39
# of foreign links	27,000	26,000	31,000	30,000
R-squared	0.388	0.437	0.389	0.438

All specifications regress log of affilates' capital intensity (defined as total assets divided by total wage bill) on home country endowment (log of capital stock divided by efficiency units of labor) and affilate-level characteristics conditional on host country × NACE 4-digit industry fixed effects. MNE refers to the sample of multinational affiliates. Standard errors are clustered at both home country and host country × industry levels. Significance levels: * 0.10 ** 0.05 *** 0.01. Number of observations is rounded to thousands of firms.

I present the distribution and coverage of the sample by country in Table OA.4. The sample covers 21 host countries and 22 home countries.[1] The coverage of the data varies across countries. I calculate the share of employment of Orbis firms in the country's total employment in Column 2. The coverage can be as high as 91% as in Bulgaria, and as low as 7% as in Poland. On average, Orbis covers 42% of the employment in each country. I also compare the total employment reported by foreign affiliates identified by Orbis to the total employment of foreign firms in the OECD/Eurostat Foreign Affiliates Statistics Database. On average, Orbis covers 21% of such employment, and the variation is smaller across host countries compared to the coverage of aggregate employment.

Table OA.4: Coverage of the firm-level data

Country	# of firms	Employment share	# of inward affiliates	Employment share	# outward affiliates
Belgium	11000	0.30	1513		1156
Bulgaria	140000	0.91	513	0.15	279
Czech	55000	0.45	2265	0.35	128
Germany	33000	0.13	2068	0.16	7171
Denmark	2000	0.08	263	0.14	715
Estonia	32000	0.51	497	0.23	25
Spain	424000	0.41	3778	0.28	1645
Finland	44000	0.34	314	0.12	653
France	207000	0.19	3043	0.20	3567
UK	42000	0.33	4945	0.36	2865
Croatia	54000	0.59	519	0.25	
Hungary	201000	0.63	350	0.16	797
Italy	417000	0.37	3659	0.23	3963
Japan	208000	0.19	69	0.04	1670
Korea	65000		401		164
Norway	86000	0.70	818	0.24	466
Poland	11000	0.07	684	0.06	
Portugal	207000	0.57	1480	0.35	405
Romania	298000	0.72	5456	0.27	205
Serbia	34000		1105		81
Slovenia	40000	0.53	414	0.25	
US	7000		0		7057
Average	119000	0.42	1552	0.21	1737

Total number of firms is rounded to 1000.

[1]The United States is the home country of around 7,000 firms in my data, but it does not show up in the data as a host country. This is because Orbis does not include any private firms operating in the United States.

My sample does not cover the universe of non-MNE companies and multinational affiliates in the host countries. In Table OA.5, I show that the technology origin and size effects are not driven by different sample coverage across countries. In Columns 1-3, I re-run the regressions in Columns 1, 2 and 4 in Table 1 of the paper but only focus on firms in host countries where my data cover at least 30% of total employment in the business sector. This leaves me with only 16 out of the 21 host countries in the original sample, but has little effect on the empirical patterns. In Column 4, I interact home country capital abundance with the coverage of my data for each host country, and the interaction term turns out to be insignificant. Therefore, the technology origin effect identified using firms in countries with poor coverage is not significantly different from that identified using firms in countries with good coverage. Finally, in Column 5, I perform a similar regression but focus on the MNE firms and replace the coverage variable using the ratio of total employment of all foreign affiliates in my data to that reported in the OECD/Eurostat databases. This ratio captures the selection bias in the multinational firm subsample. The interaction term is again insignificant.

Table OA.5: Robustness to various ways of controlling Orbis coverage

Sample firms	Orbis Coverage $\geq$ 30%			All firms	MNE
	(1)	(2)	(3)	(4)	(5)
$\log(K_i/L_i)$	0.188**		0.174**	0.289***	0.293*
	(0.0840)		(0.0679)	(0.0750)	(0.143)
$\mathbb{1}_{MNE}$		-0.0487			
		(0.0791)			
$\mathbb{1}_{MNE} \times (\log(K_i/L_i) - \log(K_l/L_l))$		0.222***			
		(0.0596)			
$\log(K_i/L_i)\times$ Orbis Coverage				-0.148	
				(0.172)	
$\log(K_i/L_i)\times$ Orbis Coverage of Inward MP Employment					-0.332
					(0.799)
log(revenue)	0.0693**	0.0699**	0.0279		0.0561***
	(0.0277)	(0.0283)	(0.0192)		(0.0104)
debt-to-equity ratio			0.00391***		
			(0.00136)		
# of firms	2,158,000	2,158,000	1,647,000	2,520,000	51,000
# of host countries	16	16	16	19	18
# of host * industry	5,830	5,830	5,787	6,779	3,631
# of home countries	22	22	22	22	22
# of foreign links	28,000	28,000	21,000	32,000	30,000
R-squared	0.333	0.333	0.364	0.337	0.436

All specifications regress log of affiliates' capital intensity (defined as total assets divided by wage bill) on home country endowment (log of capital stock divided by efficiency units of labor) and affiliate-level characteristics conditional on host country $\times$ NACE 4-digit industry fixed effects.
Orbis Coverage is measured as the employment share of Orbis firms in total business sector employment from aggregate statistics (Eurostat and OECD), for all firms (columns 1-4) and for foreign affiliates (column 5), respectively. In Table OA.4, I report these statistics for each host country. Note that I subtract the median of these statistics before interacting them with other variables so that the coefficients before home country endowments are comparable to my main specification.
MNE refers to the subsample of multinational affiliates. Standard errors are clustered at both home country and host country $\times$ industry levels. Significance levels: * 0.1 ** 0.05 *** 0.01. Number of observations is rounded to thousands of affiliates.

OA.2 Additional Empirical Results

In this section, I present additional empirical results regarding the technology origin and size effects. These two effects are robust to specifications with additional/alternative controls and to using different sub-samples.

OA.2.1 Using K_f/L_f as the Dependent Variable

In the paper, I use firms' assets-to-wage-bill ratios as measures of their capital intensities. We know from the literature that larger firms tend to pay higher wages (Oi, 1961). If the differences in wages across firms reflect heterogeneity in workers' skills, using wage bill as a denominator can capture the efficiency units of labor hired by each firm. However, this adjustment is based on the assumption that workers with different level of skills are perfect substitutes. In this subsection, I re-run the regressions in Table 1 using assets-to-employment ratios as the dependent variable to examine how the measurement of labor affects the main empirical results.

Table OA.6: Size and Technology Origin Effects on $\log(K_a/L_a)$

	Dependent Var: log(total assets/employment)					
	All Firms		MNE	All Firms		MNE
	(1)	(2)	(3)	(4)	(5)	(6)
$\log(K_i/L_i)$	0.365***		0.275***	0.315***		0.267**
	(0.0562)		(0.0897)	(0.0475)		(0.0953)
$\mathbb{1}_{MNE}$		0.0496			0.0765	
		(0.0593)			(0.0508)	
$\mathbb{1}_{MNE} \times (\log(K_i/L_i) - \log(K_l/L_l))$		0.357***			0.310***	
		(0.0469)			(0.0442)	
log(revenue)	0.276***	0.275***	0.155***	0.239***	0.238***	0.137***
	(0.0247)	(0.0254)	(0.00867)	(0.0198)	(0.0205)	(0.00981)
debt-to-equity ratio				0.00212***	0.00212***	0.00217***
				(0.000669)	(0.000669)	(0.000475)
# of firms	2,621,000	2,621,000	55,000	2,009,000	2,009,000	45,000
# of host countries	21	21	21	21	21	21
# of host * industry	7,404	7,404	4,112	7,317	7,317	3,736
# of home countries	22	22	22	22	22	22
# of foreign links	34,000	34,000	33,000	26,000	26,000	26,000
R-squared	0.497	0.497	0.440	0.479	0.479	0.445

All specifications regress log of affiliates' capital intensity (defined as total assets divided by number of employees) on home country endowment (log of capital stock divided by efficiency units of labor) and affiliate-level characteristics conditional on host country × NACE 4-digit industry fixed effects. MNE refers to the sample of multinational affiliates. Standard errors are clustered at both home country and host country × industry levels. Significance levels: * 0.10 ** 0.05 *** 0.01. Number of observations is rounded to thousands of affiliates.

In Table OA.6, I replicate the regressions in Table 1 in the paper just by replacing the dependent variable with affiliates' asset-to-employment ratios. The technology origin effects are slightly larger than those estimated in Table 1, while the sized effects are much larger than those estimated with asset-to-wage-bill ratios (0.04-0.08 in the paper). These results suggest that the elasticity of affiliate wage with respect to affiliate revenue is around 0.1 to 0.21. Not controlling for affiliate wage will greatly bias the size effects upward.

In Table OA.7, I keep using assets-to-employment ratio as the dependent variable but control affiliate wage explicitly in the regressions. Controlling affiliate wage leads to smaller size effects. The coefficients before affiliate wage is estimated between 0.33 and 0.51. Note that if we assume that affiliate in the same host-country-industry cell face the same rental rates, this coefficient can be seen as the elasticity of substitution between capital and labor. Though the elasticities here are identified under different identification assumptions (exogenous variation in affiliates' wage within a country-industry cell) from those in the paper (equation (16)), the numbers are similar.

Table OA.7: Size and Technology Origin Effects on $\log(K_a/L_a)$, Controlling $\log(w_a)$

	Dependent Var: log(total assets/employment)					
	All Firms		MNE	All Firms		MNE
	(1)	(2)	(3)	(4)	(5)	(6)
$\log(K_i/L_i)$	0.320***		0.274**	0.275***		0.275**
	(0.0593)		(0.102)	(0.0506)		(0.104)
$\mathbb{1}_{MNE}$		0.0308			0.0655	
		(0.0578)			(0.0490)	
$\mathbb{1}_{MNE} \times (\log(K_i/L_i) - \log(K_l/L_l))$		0.315***			0.270***	
		(0.0523)			(0.0474)	
log(revenue)	0.211***	0.211***	0.107***	0.172***	0.171***	0.0895***
	(0.0305)	(0.0311)	(0.00887)	(0.0235)	(0.0241)	(0.0106)
debt-to-equity ratio				0.00273***	0.00273***	0.00294***
				(0.000843)	(0.000842)	(0.000524)
$\log(w_a)$	0.330***	0.330***	0.487***	0.341***	0.341***	0.508***
	(0.0486)	(0.0485)	(0.0466)	(0.0464)	(0.0463)	(0.0488)
# of firms	2,621,000	2,621,000	55,000	2,009,000	2,009,000	45,000
# of host countries	21	21	21	21	21	21
# of host * industry	7,404	7,404	4,112	7,317	7,317	3,736
# of home countries	22	22	22	22	22	22
# of foreign links	34,000	34,000	33,000	26,000	26,000	26,000
R-squared	0.513	0.513	0.465	0.499	0.499	0.473

All specifications regress log of affiliates' capital intensity (defined as total assets divided by number of employees) on home country endowment (log of capital stock divided by efficiency units of labor) and affiliate-level characteristics conditional on host country × NACE 4-digit industry fixed effects. MNE refers to the sample of multinational affiliates. Standard errors are clustered at both home country and host country × industry levels. Significance levels: * 0.10 ** 0.05 *** 0.01. Number of observations is rounded to thousands of affiliates.

OA.2.2 Decompose Variation in Affiliates' Capital Intensities

In this subsection, I examine how much different sets of fixed effects can explain affiliates' capital-labor ratios. I focus on multinational affiliate subsample since I am interested in the explanatory power of parent-firm fixed effects. I also examine the explanatory power of the host-country-industry fixed effects, which I control in all of my reduced-form regressions.

Table OA.8 reports the results. When the host-country-industry and parent-firm fixed effects are added separately, the two sets of fixed effects produce R-squared of 0.46 and 0.41, respectively. When both are added, the R-squared increases to 0.61. There is a strong parent-firm component in affiliates' capital-labor ratios, which is consistent with our model.

Table OA.8: Decompose Variation in Affiliates' K/wL

	(1)	(2)	(3)
Host country * Industry FE	Yes	No	Yes
Parent FE	No	Yes	Yes
# of aff	55,000	55,000	55,000
R-squared	0.457	0.411	0.614

The sample is all multinational affiliates with non-missing capital intensities. The number of observations is rounded to thousands of affiliates.

OA.2.3 Impact of Parent Capital Intensity

In this subsection, I explore the impact of parent firm capital intensity on that of the affiliates. In particular, I start from regressions in Table 1 in the paper and add measures of the parent firm capital intensity. There are two caveats about these regressions. In my data, I do not observe parent firm total assets or employment for about half of the affiliates. Moreover, Orbis only provide employment rather than wage bill information for the parent firms, so the parents' capital intensity measure is slightly different from the affiliates'.[2]

With these caveats in mind, Table OA.9 shows that (1) parent capital intensity is positively correlated with that of the affiliates and (2) the technology origin and size effects are attenuated but do not disappear after controlling parent capital intensity. The second fact cannot be explained by the current model, because the current model prescribes that the parent firm capital intensity perfectly explains that of the affiliates. One possible explanation is that parent firm capital intensity captures other factors that affect the affiliates' capital-labor ratio, such as the costs of financing.

Here I provide a simple statistical model to illustrate why the technology origin and size effects are not reduced to zero when parent firm capital intensities are controlled, if parent firm capital intensities indeed capture other factors beyond technology origin and firm size. Denote the affiliate's capital intensity with y, and suppose the home country capital abundance is x_1 (I assume away the size effects for simplicity). There is another factor, x_2, which is not captured in the baseline regressions. For example, x_2 may represent the costs of financing. We do not have a measure of x_2, but we know the parent firm capital intensity x_3, which is a function of x_1 and x_2. Suppose x_3 is a linear combination of the two:

$$x_3 = x_1 + x_2.$$

Now, suppose the affiliate capital intensity is determined by x_1 and x_2 as follows

$$y = \alpha_1 x_1 + \alpha_2 x_2 + \epsilon, \tag{2}$$

where ϵ is an error term that is independent of x_1, x_2. Instead of estimating this equation, we regress y on x_1 and x_3 and obtain coefficients $\hat{\beta}_1, \hat{\beta}_3$. We must have

$$\hat{\beta}_1 \to \alpha_1 - \alpha_2, \hat{\beta}_3 \to \alpha_2.$$

Therefore, as long as $\alpha_1 > \alpha_2$, we will likely obtain a positive $\hat{\beta}_1$.

Suppose we first estimate equation (2) with only one independent variable x_1, the esti-

[2] The capital stock of the parent firm is measured as the total asset divided by the asset deflator constructed in Section OA.1.2, assuming all firms have accumulated capital for ten years.

Table OA.9: The Impact of Parent Capital Intensity

	(1)	(2)	(3)	(4)
$\log(K_i/L_i)$	0.322**	0.235*	0.295**	0.207
	(0.121)	(0.115)	(0.126)	(0.122)
log(revenue)	0.0828***	0.0812***	0.0769***	0.0756***
	(0.0111)	(0.0108)	(0.0112)	(0.0108)
log(parent K/L)		0.0866***		0.0889***
		(0.0151)		(0.0140)
debt-to-equity ratio			0.00290***	0.00291**
			(0.00101)	(0.00103)
# of firms	27,000	27,000	22,000	22,000
# of host countries	21	21	21	21
# of host * industry	2,681	2,681	2,415	2,415
# of home countries	22	22	22	22
# of foreign links	14,000	14,000	12,000	12,000
R-squared	0.437	0.443	0.453	0.460

All specifications regress log of affiliates' capital intensity (defined as total assets divided by total wage bill) on home country endowment (log of capital stock divided by efficiency units of labor) and affiliate-level characteristics conditional on host country × NACE 4-digit industry fixed effects. MNE refers to the sample of multinational affiliates. Standard errors are clustered at both home country and host country × industry levels. Significance levels: * 0.10 ** 0.05 *** 0.01. Number of observations is rounded to thousands of affiliates.

mated $\hat{\alpha}_1$ will be biased upward if x_1 and x_2 are positively correlated and biased downward vice versa. In both cases, it is possible that

$$plim\,\hat{\alpha}_1 > plim\,\hat{\beta}_1 = \alpha_1 - \alpha_2.$$

OA.2.4 MNE Premium at Home

In Table OA.10, I examine whether multinational affiliates in the home country have different capital intensities compared to non-multinationals. In Columns 1 and 2, I focus on the sample of all MNE affiliates at home and non-MNEs. Therefore, there is no foreign affiliates in this sample (# of foreign links is zero), and the technology origin effects cannot be identified due to the host-country-industry fixed effects. I regress the affiliates' capital-labor ratios on a MNE dummy and affiliate revenue. MNE affiliates seem to have a positive premium in capital intensities compared to local firms, but the difference is not significant. In Column 2, I interact this term with a dummy indicating the home country's capital abundance being above the median. The interaction term is not significant, suggesting there is not much heterogeneity between capital-scarce and capital-abundant home countries. In Columns 3 and 4, I use the entire sample by including the foreign affiliates. In this case, I can identify the MNE-at-home dummy and the technology origin effects at the same time. The patterns are similar to those in Columns 1 and 2.

Table OA.10: MNE Premium at Home

	Dependent Var: log(total assets/wage bill)			
	non-MNEs + MNE Home Subsidiaries		All Firms	
	(1)	(2)	(3)	(4)
MNE affiliate at Home	0.0706	0.0529	0.0729	0.0476
	(0.0470)	(0.0519)	(0.0433)	(0.0480)
MNE affiliate at Home $\times \mathbb{1}(K_i/L_i \geq \text{median})$		0.0250		0.0357
		(0.0465)		(0.0468)
$\log(K_i/L_i)$			0.173*	0.173*
			(0.0844)	(0.0847)
log(revenue)	0.0814***	0.0814***	0.0799***	0.0799***
	(0.0276)	(0.0276)	(0.0268)	(0.0267)
# of firms	2,585,000	2,585,000	2,621,000	2,621,000
# of host countries	21	21	21	21
# of host * industry	7,356	7,356	7,404	7,404
# of home countries	21	21	22	22
# of foreign links	0	0	34,000	34,000
R-squared	0.357	0.357	0.356	0.356

All specifications regress log of affiliates' capital intensity (defined as total assets divided by total wage bill) on home country endowment (log of capital stock divided by efficiency units of labor) and affiliate-level characteristics conditional on host country × NACE 4-digit industry fixed effects. MNE refers to the sample of multinational affiliates. Standard errors are clustered at both home country and host country × industry levels. Significance levels: * 0.10 ** 0.05 *** 0.01. Number of observations is rounded to thousands of affiliates.

OA.2.5 Controlling Gravity

In this subsection, I perform a robustness check of the technology origin and size effects by adding gravity controls.[3] The worry is that MNCs may compensate the distance between the affiliates and the home country with the addition of capital embedded knowledge, a mechanism related to Keller and Yeaple (2013).

I report the results in Table OA.11. Columns 1-4 use all firms while columns 5-8 focus on the multinational affiliate subsample. The gravity variables controlled include log distance between the home and host countries and dummies that indicate whether the home and host countries share a border, use a common language or have a colonial relationship. Most of the gravity variables are insignificant. The only exception is the contiguity dummy, which is negative and significant in Columns 7 and 8. The negative effect is somewhat consistent with the capital embedded knowledge mechanism. However, both the technology origin and size effects are robust.

Table OA.11: Size and Technology Origin Effects on $\log(K_a/w_a L_a)$, Controlling Gravity

	Dependent Var: log(total assets/wage bill)							
	All Firms				MNE			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\log(K_i/L_i)$	0.237***	0.197***	0.232***	0.197***	0.252**	0.258**	0.238*	0.251**
	(0.0626)	(0.0538)	(0.0637)	(0.0547)	(0.121)	(0.123)	(0.121)	(0.120)
log(revenue)	0.0810***	0.0432*	0.0811***	0.0433*	0.0557***	0.0439***	0.0547***	0.0429***
	(0.0267)	(0.0218)	(0.0267)	(0.0219)	(0.00996)	(0.0107)	(0.0100)	(0.0108)
debt-to-equity ratio		0.00390***		0.00390***		0.00367***		0.00371***
		(0.00125)		(0.00125)		(0.000638)		(0.000647)
log(distance)	-0.0243	-0.00657	-0.0201	-0.00716	-0.0238	-0.0224	-0.0255	-0.0245
	(0.0247)	(0.0212)	(0.0256)	(0.0217)	(0.0154)	(0.0151)	(0.0187)	(0.0178)
Contiguity			-0.0744	-0.0330			-0.110***	-0.105***
			(0.0642)	(0.0605)			(0.0302)	(0.0269)
Common language			0.0116	-0.0467			-0.0497	-0.0684
			(0.0846)	(0.0798)			(0.0460)	(0.0458)
Colonial relationship			-0.0270	0.0718			0.0126	0.0455
			(0.0514)	(0.0590)			(0.0511)	(0.0472)
# of firms	2,621,000	2,009,000	2,621,000	2,009,000	55,000	45,000	55,000	45,000
# of host countries	21	21	21	21	21	21	21	21
# of host * industry	7,404	7,317	7,404	7,317	4,112	3,736	4,112	3,736
# of home countries	22	22	22	22	22	22	22	22
# of foreign links	34,000	26,000	34,000	26,000	33,000	26,000	33,000	26,000
R-squared	0.380	0.421	0.380	0.421	0.463	0.484	0.464	0.484

All specifications regress log of affiliates' capital intensity (defined as total assets divided by total wage bill) on home country endowment (log of capital stock divided by efficiency units of labor) and affiliate-level characteristics conditional on host country × NACE 4-digit industry fixed effects. MNE refers to the sample of multinational affiliates. Standard errors are clustered at both home country and host country × industry levels. Significance levels: * 0.10 ** 0.05 *** 0.01. Number of observations is rounded to thousands of affiliates.

[3] I thank a referee for suggesting this robustness check.

OA.2.6 Controlling Revenue and Age Categories

In Table OA.12, I perform a robustness check of the technology origin effect by controlling the size effect in a more flexible way.[4] Instead of controlling a linear term of affiliate revenue, I control for a set of size-bin-industry-location fixed effects. For each affiliate in a particular industry-location cell, I group it into a particular size bin based on its revenue (either six or ten equally-sized bins). In Columns 3 and 6, I further control for a set of age-bin-industry-location fixed effects, where I use the date of incorporation in my data to define affiliate age. The technology origin effect is robust to these alternative controls.

Table OA.12: Technology Origin Effects: Controlling Revenue and Age Categories

	Dependent Var: log(total assets/wage bill)					
	All Firms			MNE		
	(1)	(2)	(3)	(4)	(5)	(6)
$\log(K_i/L_i)$	0.199***	0.212***	0.214***	0.284**	0.296**	0.338***
	(0.0599)	(0.0602)	(0.0665)	(0.111)	(0.112)	(0.110)
debt-to-equity ratio	0.00407***	0.00406***	0.00542***	0.00312***	0.00323***	0.00315***
	(0.00106)	(0.00105)	(0.00112)	(0.000636)	(0.000693)	(0.00102)
# of firms	2,009,000	2,009,000	1,937,000	45,000	45,000	45,000
# of host countries	21	21	21	21	21	21
# of host * industry	6,702	6,315	5,797	3,079	2,755	1,682
# of home countries	22	22	22	22	22	22
# of foreign links	26,000	26,000	25,000	23,000	21,000	16,000
R-squared	0.449	0.457	0.485	0.559	0.579	0.656
Revenue Categories	6 Bins	10 Bins	10 Bins	6 Bins	10 Bins	10 Bins
Age Categories	NA	NA	10 Bins	NA	NA	10 Bins

All specifications regress log of firms' capital intensity (defined as total assets divided by total wage bill) on home country endowment (log of capital stock divided by efficiency units of labor) and firm-level characteristics conditional on host country × NACE 4-digit industry × revenue category fixed effects (and host × industry × age category fixed effects). MNE refers to the sample of multinational subsidiaries. Standard errors are clustered at both home country and host country × industry levels. Significance levels: * 0.10 ** 0.05 *** 0.01. Number of observations is rounded to thousands of firms. The number of observations shrinks when finer fixed effects are added since singletons are dropped.

[4] I thank a referee for suggesting this robustness check.

OA.2.7 Regressions at Firm-Sector-Location Level

In this subsection, I perform a robustness check based on data aggregated at firm-sector-location level, where sector refers to four-digit NACE Rev 2 industries and location refers to the production (host) country. For observations in the original data that are non-MNE companies or single affiliates of a multinational in a particular sector-location cell, this aggregation has no impact on the dependent and independent variables. For parent firms that have multiple affiliates in a sector-location cell, I add up the affiliates' total assets, wage bills and revenue and use them to calculate the dependent and independent variables. The debt-to-equity ratio is also re-calculated using the aggregated affiliate-level debt and equity. Since it is rare to see multiple affiliates in a sector-location cell, this aggregation only reduces the sample size slightly. The results are similar to those in the main specification (see Table 1 in the paper).

Table OA.13: Size and Technology Origin Effects: Regressions at Firm-Section-Location Level

	Dependent Var: log(total assets/wage bill)			
	All Firms		MNE	
	(1)	(2)	(3)	(4)
$\log(K_i/L_i)$	0.220***	0.190***	0.265**	0.281**
	(0.0670)	(0.0597)	(0.123)	(0.122)
log(revenue)	0.0811***	0.0434*	0.0523***	0.0423***
	(0.0266)	(0.0216)	(0.0102)	(0.0112)
debt-to-equity ratio		0.00395***		0.00384***
		(0.00129)		(0.000735)
# of firm-host-industry	2,580,000	1,976,000	45,000	37,000
# of host countries	21	21	21	21
# of host * industry	7,401	7,315	4,019	3,647
# of home countries	22	22	22	22
R-squared	0.315	0.344	0.390	0.415

All specifications regress log of capital intensity (defined as total assets divided by total wage bill, aggregated at firm-host-industry level) on home country endowment (log of capital stock divided by efficiency units of labor) and firm-host-industry level characteristics conditional on host country × NACE 4-digit industry fixed effects. MNE refers to the sample of multinational firms. Standard errors are clustered at both home country and host country × industry levels. Significance levels: * 0.10 ** 0.05 *** 0.01. Number of observations is rounded to thousands of firm-host-country-industry combinations.

OA.2.8 Horizontal MP

Table OA.14 reports regressions where I interact the home country capital abundance with a dummy that indicates whether the affiliate can be seen as a "horizontal" foreign investment by the parent firm. Affiliates are defined as "horizontal" if they are in the same two-, three- or four-digit NACE Rev.2 industry as their parent firms. The results suggest that whether the affiliate is "horizontal" or not, home country capital abundance predicts a higher capital-labor ratio. Though insignificant, the technology origin effect is 20-40% larger for horizontal affiliates, with a larger coefficient when I use finer industry classifications to define "horizontal".

The fact that the technology origin effects also show up among non-horizontal affiliates is consistent with the knowledge capital view of the multinational firms (Ethier and Markusen, 1996), which is also adopted in this paper. Several papers have shown that the majority of domestic and foreign affiliates in the upstream industries of the parent firms do not ship their output for further use by the parents. (Atalay et al., 2014; Ramondo et al., 2016) One explanation is that the parent firms may transfer their technology know-hows even to non-horizontal affiliates, which can potentially explain the observed technology origin effects for all affiliates. Consistent with this view, I do not distinguish horizontal and other types of MP in the main regressions and the quantitative implementations in the paper.

Table OA.14: Interacting Technology Origin with Horizontal MP Indicator

	Dependent Var: log(total assets/wage bill)					
Definition of horizontal MP:	Same NACE 2-digit		Same NACE 3-digit		Same NACE 4-digit	
	(1)	(2)	(3)	(4)	(5)	(6)
$\log(K_i/L_i)$	0.265**	0.274**	0.263**	0.270**	0.263**	0.270**
	(0.127)	(0.127)	(0.124)	(0.123)	(0.122)	(0.122)
$\log(K_i/L_i)\times$ horizontal	0.0531	0.0541	0.0803	0.0956	0.0948	0.112
	(0.114)	(0.112)	(0.108)	(0.0987)	(0.106)	(0.0980)
horizontal	-0.540	-0.547	-0.838	-1.016	-0.991	-1.191
	(1.319)	(1.293)	(1.253)	(1.144)	(1.228)	(1.132)
log(revenue)	0.0544***	0.0425***	0.0545***	0.0427***	0.0546***	0.0429***
	(0.00982)	(0.0106)	(0.00987)	(0.0107)	(0.00986)	(0.0106)
debt-to-equity ratio		0.00372***		0.00372***		0.00373***
		(0.000642)		(0.000644)		(0.000647)
# of firms	55,000	45,000	55,000	45,000	55,000	45,000
# of host countries	21	21	21	21	21	21
# of host * industry	4,112	3,736	4,112	3,736	4,112	3,736
# of home countries	22	22	22	22	22	22
# of foreign links	33,000	26,000	33,000	26,000	33,000	26,000
# of horizontal links	9,774	8,191	7,505	6,285	6,354	5,328
# of vertical links	7,139	5,963	7,139	5,963	7,139	5,963
R-squared	0.463	0.484	0.463	0.484	0.463	0.484

All specifications regress log of affiliates' capital intensity (defined as total assets divided by total wage bill, aggregated at firm-host-industry level) on home country endowment (log of capital stock divided by efficiency units of labor) and affiliate-host-industry level characteristics conditional on host country × NACE 4-digit industry fixed effects. MNE refers to the sample of multinational affiliates. Standard errors are clustered at both home country and host country × industry levels. Significance levels: * 0.10 ** 0.05 *** 0.01. Number of observations is rounded to thousands of affiliates.

OA.2.9 Further Controlling Heterogeneous Financing Costs

In this subsection, I provide two additional robustness checks regarding heterogenous financing costs across affiliates. In Table OA.15, I focus on the role of the "global debt-to-equity ratio" of the firm instead of the affiliate-level debt-to-equity ratio. The global ratio is calculated using total debt and total equity of all affiliates that belong to the same parent firm. Therefore, it captures the availability of funds of the entire multinational corporation.[5] Controlling for global debt-to-equity ratio barely changes the estimated technology origin and size effects. When controlling for both global and affiliate-level debt-to-equity ratio (Column 3 of Table OA.15), both controls are significant, and the coefficient of the global ratio is higher than the affiliate-level ratio. However, the magnitudes of the two key coefficients are similar to those in Table 1 of the paper.

Table OA.15: Size and Technology Origin Effects: Controlling Global Debt-to-Equity Ratio

	Dependent Var: log(total assets/wage bill)		
	All Firms	MNE	
	(1)	(2)	(3)
$\log(K_i/L_i)$	0.191***	0.295**	0.286**
	(0.0499)	(0.111)	(0.117)
log(revenue)	0.0430*	0.0515***	0.0446***
	(0.0214)	(0.0106)	(0.0110)
global debt-to-equity	0.00382**	0.00571***	0.00440**
	(0.00139)	(0.00172)	(0.00201)
debt-to-equity ratio			0.00247***
			(0.000617)
# of firms	2,039,000	52,000	45,000
# of host countries	21	21	21
# of host * industry	7,330	3,994	3,736
# of home countries	22	22	22
# of foreign links	30,000	29,000	26,000
R-squared	0.420	0.471	0.484

All specifications regress log of affiliates' capital intensity (defined as total assets divided by total wage bill, aggregated at firm-host-industry level) on home country endowment (log of capital stock divided by efficiency units of labor) and affiliate-level and firm-level characteristics conditional on host country × NACE 4-digit industry fixed effects. MNE refers to the sample of multinational affiliates. Standard errors are clustered at both home country and host country × industry levels. Significance levels: * 0.10 ** 0.05 *** 0.01. Number of observations is rounded to thousands of affiliates.

In Table OA.16, I consider controlling affiliate age and debt-to-equity ratios (leverage ratios) using age-bin-industry-location and leverage-bin-industry-location fixed effects. I first group firms within each industry-location cell into 10 equally-sized bins based on their age and leverage ratios. I then add age-bin-industry-location and leverage-bin-industry-location

[5] Note that the global ratio is only different from the affiliate-level ratio when the affiliate belongs to a multinational corporation and has other sibling affiliates that belong to the same parent firm in my dataset.

fixed effects in the regressions. These fixed effects control the impact of affiliate age and leverage ratios in a flexible way. Affiliate age may have an impact of external financing since it takes time for them to accumulate net worth that can be used as collateral. In all specifications, the technology origin and size effects are significant and stable, expect for Column 3 where I control both sets of fixed effects for the sample of all firms – the size effect becomes smaller and insignificant. When I control both sets of fixed effects for the multinational affiliate subsample, the technology origin effect is even slightly larger than that in Column 3 of Table 1 in the paper. In Column 7 of Table OA.16, I further control for the global-leverage-bin-industry-location fixed effects to capture the availability of funds of the entire multinational corporation (a similar idea to Column 3 of Table OA.15). The results are robust.

Table OA.16: Size Technology Origin Effects: Controlling Leverage Ratio and Age Categories

	Dependent Var: log(total assets/wage bill)						
	All Firms			MNE			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\log(K_i/L_i)$	0.243***	0.224***	0.228***	0.310***	0.308**	0.352***	0.369***
	(0.0688)	(0.0566)	(0.0588)	(0.107)	(0.111)	(0.112)	(0.114)
log(revenue)	0.0630**	0.0500**	0.0291	0.0567***	0.0595***	0.0606***	0.0569***
	(0.0295)	(0.0227)	(0.0238)	(0.0111)	(0.0130)	(0.0138)	(0.0142)
# of firms	2,524,000	2,009,000	1,937,000	55,000	45,000	45,000	45,000
# of host countries	21	21	21	21	21	21	20
# of host * industry	6,777	6,152	5,660	2,695	2,225	1,252	812
# of home countries	22	22	22	22	22	22	22
# of foreign links	33,000	26,000	25,000	26,000	19,000	15,000	13,000
R-squared	0.409	0.450	0.478	0.561	0.590	0.666	0.723
Age Bins	Yes	No	Yes	Yes	No	Yes	Yes
Leverage Bins	No	Yes	Yes	No	Yes	Yes	Yes
Global Leverage Bins	No	No	No	No	No	No	Yes

All specifications regress log of affiliates' capital intensity (defined as total assets divided by total wage bill) on home country endowment (log of capital stock divided by efficiency units of labor) and affiliate-level characteristics conditional on host country × NACE 4-digit industry × 10 age bins fixed effects (and host × industry × 10 leverage ratio bins fixed effects). MNE refers to the sample of multinational affiliates. Standard errors are clustered at both home country and host country × industry levels. Significance levels: * 0.10 ** 0.05 *** 0.01. Number of observations is rounded to thousands of affiliates.

OA.2.10 Focus on Subsample with Debt-to-equity Ratios

In this subsection, we replicate Table 1 of the paper by focusing on the subsample of affiliates with non-missing debt-to-equity ratios. This does not change Columns 4-6 of the table since debt-to-equity ratio is explicitly controlled there. For Columns 1-3, the results are close to those in Columns 4-6, and similar to those in Columns 1-3 of Table 1 of the paper.

Table OA.17: Size and Technology Origin Effects on $\log(K_a/w_a L_a)$

	Dependent Var: log(total assets/wage bill)					
	All Firms		MNE	All Firms		MNE
	(1)	(2)	(3)	(4)	(5)	(6)
$\log(K_i/L_i)$	0.160**		0.287*	0.158**		0.291*
	(0.0707)		(0.146)	(0.0700)		(0.144)
$\mathbb{1}_{MNE}$		0.0311			0.0309	
		(0.0573)			(0.0576)	
$\mathbb{1}_{MNE} \times (\log(K_i/L_i) - \log(K_l/L_l))$		0.139**			0.138**	
		(0.0539)			(0.0531)	
log(revenue)	0.0429*	0.0423*	0.0420***	0.0428*	0.0423*	0.0422***
	(0.0216)	(0.0222)	(0.0107)	(0.0216)	(0.0223)	(0.0106)
debt-to-equity ratio				0.00390***	0.00390***	0.00379***
				(0.00125)	(0.00125)	(0.000640)
# of firms	2,009,000	2,009,000	45,000	2,009,000	2,009,000	45,000
# of host countries	21	21	21	21	21	21
# of host * industry	7,317	7,317	3,736	7,317	7,317	3,736
# of home countries	22	22	22	22	22	22
# of foreign links	26,000	26,000	26,000	26,000	26,000	26,000
R-squared	0.395	0.395	0.458	0.395	0.395	0.459

All specifications regress log of affiliates' capital intensity (defined as total assets divided by total wage bill) on home country endowment (log of capital stock divided by efficiency units of labor) and affiliate-level characteristics conditional on host country × NACE 4-digit industry fixed effects. MNE refers to the sample of multinational affiliates. Standard errors are clustered at both home country and host country × industry levels. Significance levels: * 0.10 ** 0.05 *** 0.01. Number of observations is rounded to thousands of affiliates.

OA.2.11 Variation Across/Within North and South

To explore the variation across and within the group of developed and less-developed countries, I first divide the home countries into two subsamples: the "North" (capital abundance greater than median) and the "South" (capital abundance below the median). The average aggregate capital-labor ratio of Northern countries is 93% higher than that of the average ratio of Southern countries. The country with the lowest capital abundance is Bulgaria (139% lower than the average Southern country). There is also variation in aggregate capital-labor ratios within each subsample. The standard deviation is 23 log points among Northern countries and 37 log points among Southern countries.

In Table OA.18, I explore the variation in home country K_i/L_i both within and between the two groups. However, because I only have 22 home countries (11 in each group), the technology origin effects are likely to be imprecisely estimated and should be taken with caveats. In Column 1, I regress the affiliates' capital-labor ratio on a dummy variable indicating the home country being in the North, as well as interaction terms between the North dummy and affiliate revenue. The estimates suggest a large but imprecise effect of North dummy (the technology origin effect), and the size effect seems to be slightly larger among Southern firms than among Northern firms.

In Column 2, I further interact the North dummy with the affiliate's home country aggregate capital-labor ratio (normalized by the affiliate's host country capital-labor ratio to facilitate interpretation). It seems that within each group, home country capital abundance is still positively correlated with affiliates' capital-labor ratio, but with only 11 home countries in each group, it is hard to conclude whether the overall technology origin effect comes from between- or within-group variation. In the paper, I estimate the effects using the variation in capital abundance across all home countries and match the quantitative model to those coefficients.

Table OA.18: North Dummy and Interaction Terms

	Dependent Var: log(total assets/wage bill)		
	(1)	(2)	(3)
$\mathbb{1}(K_i/L_i \geq \text{median})$	0.621*	0.205	0.151
	(0.331)	(0.274)	(0.332)
$\mathbb{1}(K_i/L_i < \text{median}) \times (\log(K_i/L_i) - \log(K_l/L_l))$		0.242	0.283**
		(0.150)	(0.133)
$\mathbb{1}(K_i/L_i \geq \text{median}) \times (\log(K_i/L_i) - \log(K_l/L_l))$		0.441***	0.478***
		(0.115)	(0.109)
$\mathbb{1}(K_i/L_i < \text{median}) \times \log(\text{revenue})$	0.0865***	0.0739***	0.0602***
	(0.0146)	(0.0143)	(0.0202)
$\mathbb{1}(K_i/L_i \geq \text{median}) \times \log(\text{revenue})$	0.0448***	0.0508***	0.0399***
	(0.0119)	(0.0108)	(0.0114)
$\mathbb{1}(K_i/L_i < \text{median}) \times$ debt-to-equity			0.00523**
			(0.00236)
$\mathbb{1}(K_i/L_i \geq \text{median}) \times$ debt-to-equity			0.00297***
			(0.000434)
# of firms	55,000	55,000	45,000
# of host countries	21	21	21
# of host * industry	4,112	4,112	3,736
# of home countries	22	22	22
# of foreign links	33,000	33,000	26,000
R-squared	0.461	0.465	0.486

All specifications regress log of affiliates' capital intensity (defined as total assets divided by total wage bill, aggregated at firm-host-industry level) on home country endowment (log of capital stock divided by efficiency units of labor) and affiliate-level characteristics conditional on host country × NACE 4-digit industry fixed effects. Standard errors are clustered at both home country and host country × industry levels. Significance levels: * 0.10 ** 0.05 *** 0.01. The sample here contains only multinational affiliate. The number of observations is rounded to thousands of affiliates.

OA.2.12 Alternative Definitions of Technology Origins

In this subsection, I examine the robustness of the technology origin effects when I use three alternative definitions of "technology origins". Orbis provides information on "controlling shareholders" at each layer of the ownership pyramid. Using the information on the ownership pyramid, I define technology origins of MNE affiliates as follows:

Def 1: The country of a foreign controlling shareholder that is (1) an industrial company (2) closest to the affiliate and within two layers of control (3) in the same 2-digit industry as the affiliate.

Def 2: The country of a foreign controlling shareholder that is (1) an industrial company (2) closest to the affiliate and within three layers of control.

Def 3: The country of a foreign controlling shareholder that is (1) an industrial company (2) closest to the affiliate and within four layers of control.

I then replicate the regressions in Columns 3 and 6 of Table 1 of the paper using these alternative definitions. The results are shown in Table OA.19. The results are similar to those when I use the global ultimate owner's country as "technology origins".

Table OA.19: Alternative Definitions of Technology Origin

	Alter Def 1		Alter Def 2		Alter Def 3	
	(1)	(2)	(3)	(4)	(5)	(6)
$log(K_i/L_i)$	0.140*	0.175*	0.233	0.239	0.263*	0.272*
	(0.0789)	(0.0904)	(0.140)	(0.151)	(0.135)	(0.144)
log(revenue)	0.0800***	0.0699***	0.0695***	0.0617***	0.0692***	0.0616***
	(0.0130)	(0.0130)	(0.00893)	(0.0106)	(0.00935)	(0.0108)
debt-to-equity ratio		0.00298**		0.00322***		0.00325***
		(0.00112)		(0.000702)		(0.000622)
# of firms	27,000	22,000	44,000	36,000	44,000	37,000
# of host countries	21	21	21	21	21	21
# of host * industry	2,772	2,490	3,632	3,263	3,634	3,268
# of home countries	40	39	47	45	47	45
# of foreign links	15,000	12,000	26,000	21,000	26,000	21,000
R-squared	0.433	0.453	0.427	0.443	0.427	0.443

Two-way clustered standard errors in parentheses. * 0.10 ** 0.05 *** 0.01. Subscript i indicates the home country of the firm. Each regression controls host country × NACE 4-digit industry fixed effects.

Alternative definition 1: Foreign owners' countries are defined as technology origin only if the owner is the closest to the affiliate within 2 layers of control. The owner has to be an industrial firm and must be within the same 2-digit industry as the affiliate.

Alternative definition 2: Foreign owners' countries are defined as technology origin only if the owner is the closest to the affiliate within 3 layers of control and is an industrial firm.

Alternative definition 3: Foreign owners' countries are defined as technology origin only if the owner is the closest to the affiliate within 4 layers of control and is an industrial firm.

OA.3 Theory

OA.3.1 Derive Formulas in the Model Section

In this section, I derive formulas when solving the firm's problem backwards. For Stage 3 and Stage 2, the solution is almost the same as Arkolakis et al. (2018) (hereinafter ARRY). The only difference is that all the expressions are conditional on the core productivity ϕ due to technology-capital complementarity.

Solving the Firm's Problem in Stage 3

In Stage 3, the firm knows the realization of all shocks $(\phi, \mathbf{z})$. It has chosen (a, b) in Stage 1 and has chosen which markets to serve in Stage 2. The problem in Stage 3 is to determine where to produce and from which production location to serve a certain destination market. Without loss of generality, consider the problem of choosing production location for market n. Since there is no cost of setting up a plant, the firm simply chooses the minimum cost location l

$$l = \arg\min_{m} C_{imn}\left(\phi, \mathbf{z}, a, b\right),$$

where

$$C_{imn}\left(\phi, \mathbf{z}, a, b\right) = z_{im}^{-1}\gamma_{im}\tau_{mn}\left(\lambda\left(\frac{r_m}{a}\phi^{\xi/2-1}\right)^{1-\varepsilon} + (1-\lambda)\left(\frac{w_m}{b}\phi^{-\xi/2-1}\right)^{1-\varepsilon}\right)^{1/(1-\varepsilon)}.$$

Solving the Firm's Problem in Stage 2

In Stage 2, the firm knows ϕ and has chosen (a, b). It has to predict the expected operating profit from serving each destination market and decide whether to enter or not. The operating profit is uncertain because of the country-specific productivities $\mathbf{z}$. The firm takes expectation over the joint distribution of $\mathbf{z}$. To compute the expected operating profit in market n, consider the distribution of minimum cost $C_{i\cdot n}\left(\phi, \mathbf{z}, a, b\right) \equiv \min_m C_{imn}\left(\phi, \mathbf{z}, a, b\right)$

$$\begin{aligned}
\Pr\left(C_{i\cdot n}\left(\phi, \mathbf{z}, a, b\right) \leq c\right) &= 1 - \Pr\left(C_{i\cdot n}\left(\phi, \mathbf{z}, a, b\right) > c\right) \\
&= 1 - \Pr\left(z_{il} < \frac{\zeta_{iln}\left(\phi\right)}{\phi c}, \forall l\right) \\
&= 1 - \exp\left(-\phi^{\chi}\sum_{l} T_{il}\zeta_{iln}\left(\phi\right)^{-\chi} c^{\chi}\right),
\end{aligned}$$

where $\zeta_{iln}\left(\phi\right)$ is the marginal cost for affiliates with unit Hicks-neutral productivity ($z_{il} =$

$\phi = 1$) by

$$\zeta_{iln}(\phi) \equiv \gamma_{il}\tau_{ln}\left(\lambda\left(\frac{r_l}{a}\phi^{\xi/2}\right)^{1-\varepsilon} + (1-\lambda)\left(\frac{w_l}{b}\phi^{-\xi/2}\right)^{1-\varepsilon}\right)^{1/(1-\varepsilon)}.$$

Note that I have omitted the dependency of ζ_{iln} on (a, b) for notational simplicity. Since there is no cost of setting up a plant, to compute the expected operating profit in market n, it is sufficient to know the distribution of the marginal cost $C_{i\cdot n}(\phi, \mathbf{z}, a, b)$. The expected operating profit can be integrated over this distribution.

Define

$$\Psi_{in}(\phi) \equiv \sum_l T_{il}\zeta_{iln}(\phi)^{-\chi},$$

and the expected operating profit from market n (conditional on entry) is

$$\begin{aligned} \pi_{i\cdot n}(\phi) &= \frac{X_n}{\sigma P_n^{1-\sigma}} E_{\mathbf{Z}}\left[(\tilde{\sigma} C_{i\cdot n}(\phi))^{1-\sigma}\right] \qquad (3) \\ &= \frac{\tilde{\sigma}^{1-\sigma} X_n}{\sigma P_n^{1-\sigma}} \int_0^\infty c^{1-\sigma}\phi^\chi \Psi_{in}(\phi)\chi c^{\chi-1}\exp(-\phi^\chi \Psi_{in}(\phi) c^\chi)\, dc \\ &= \frac{\tilde{\sigma}^{1-\sigma} X_n}{\sigma P_n^{1-\sigma}} \Gamma\left(\frac{\chi - \sigma + 1}{\chi}\right)\phi^{\sigma-1}\Psi_{in}(\phi)^{\frac{\sigma-1}{\chi}}, \end{aligned}$$

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

The firm enters market n as long as the expected operating profit is larger than the marketing costs $P_n F$

$$\pi_{i\cdot n}(\phi) \geq P_n F. \qquad (4)$$

When $\xi < 2$, operating profit strictly increases with ϕ and there is a cutoff productivity ϕ^*_{in} such that

$$\pi_{i\cdot n}(\phi) \geq P_n F \text{ iff } \phi \geq \phi^*_{in}.$$

Denote the indicator function

$$S_{in}(\phi) = \mathbf{1}\left[\pi_{i\cdot n}(\phi) \geq P_n F\right],$$

and we can express the expected global profit (conditional on ϕ) as

$$\pi_i(\phi) \equiv \sum_n S_{in}(\phi)\left(\pi_{i\cdot n}(\phi) - P_n F\right).$$

Solving the Firm's Problem in Stage 1

In Stage 1, the firm chooses (a, b) to maximize global expected profit $E_\phi [\pi_i (\phi)]$:

$$(a^*, b^*) = \arg \max_{(a,b)\in\Theta} E_\phi [\pi_i (\phi)]$$

where Θ is the technology menu $\{(a, b) | \theta (a, b) \leq 1\}$. Since $\theta (a, b)$ and the firm's profit strictly increase with a and b, the firm always adopts technologies on the technology frontier thus $\theta (a^*, b^*) = 1$. Define $\delta \equiv (\varepsilon - 1) \log (a/b)$ when $\varepsilon \neq 1$. Knowing δ, one can simply back out (a, b) from the condition $\theta (a, b) = 1$. Therefore, the firm's problem is equivalent to

$$\delta^* = \arg \max_{\delta\in(-\infty,\infty)} E_\phi [\pi_i (\phi, a (\delta) , b (\delta))]$$

When solving the general equilibrium, I first combine the first order condition of δ^* as one of the equilibrium conditions and check the global optimality after the equilibrium has been found. Though I cannot provide the conditions for the uniqueness of δ^*, the first-order condition approach is always sufficient to find the global maximum in practice, and I do not find any multiple equilibria using different initial guesses. The expression for the FOC under the CES functional form of $\theta (a, b)$ is

$$\frac{\partial}{\partial\delta} E_\phi [\pi_i (\phi, a (\delta) , b (\delta))] = \frac{\sigma - 1}{1 - \varepsilon} E_\phi \left[\sum_n \left(t (\delta) - \sum_l \psi_{iln} (\delta) \kappa_l (\delta) \right) S_{in} (\phi) \pi_{i\cdot n} (\phi) \right],$$

where $t (\delta) \equiv \left(1 + \exp \left(\frac{1-\eta}{1-\varepsilon}\delta\right)\right)^{-1}$. Two things are worth noting when deriving the first order conditions. First, $\pi_{i\cdot n}$ and S_{in} depend on δ. Second, a small change in δ affects the expected global profit not only by changing $\pi_{i\cdot n} (\phi)$ but also by changing S_{in}, the decision of entering markets (or equivalently, the cutoff productivity ϕ^*_{in}). However, since at the cutoff productivity ϕ^*_{in}, the operating profit equals the marketing costs, this margin adds nothing to the FOC. This result holds even when we take into account that ϕ^*_{in} might be smaller than $\phi_{\min}$, the lower bound of ϕ. One can verify this by considering the case where $\phi^*_{in} (\delta) = \phi_{\min}$ and $\phi^*_{in} (\delta + \Delta\delta) > \phi_{\min}$.

Firm's Entry

Firms headquarter in country i until the expected profit with the optimal technology equals entry costs

$$\begin{aligned} E_\phi [\pi_i (\phi, a (\delta^*) , b (\delta^*))] &\leq P_i F_{ei}, \\ E_\phi [\pi_i (\phi, a (\delta^*) , b (\delta^*))] &= P_i F_{ei} \text{ if } M_i > 0. \end{aligned}$$

OA.3.2 Aggregation and Equilibrium

In this subsection, I derive expressions for aggregate variables and define the general equilibrium of the model. The expressions are useful both for the calibration and for deriving analytical results in the paper.

Aggregate variables are expressed in integrals of firm level variables over the distribution of core productivity ϕ. Conditional on ϕ and the firm entering market n, the probability that country l becomes the lowest cost production location is

$$\psi_{iln}(\phi, a_i, b_i) \equiv \frac{T_{il}(\gamma_{il} C_l(\phi, 1, a_i, b_i) \tau_{ln})^{-\chi}}{\sum_m T_{im}(\gamma_{im} C_m(\phi, 1, a_i, b_i) \tau_{mn})^{-\chi}}, \tag{5}$$

where (a_i, b_i) is the equilibrium technology choice for all firms from country i. The expected sales from country l to n by affiliates owned by country i firms are

$$X_{iln}(\phi, a_i, b_i) = \sigma \psi_{iln}(\phi, a_i, b_i) \pi_{i \cdot n}(\phi, a_i, b_i). \tag{6}$$

To obtain aggregate sales to destination n by affiliates in country l from home country i to destination n, I integrate over all country i firms

$$X_{iln} = M_i \int S_{in}(\phi, a_i, b_i) X_{iln}(\phi, a_i, b_i) dF(\phi), \tag{7}$$

where M_i is the mass of firms headquartered in country i. Similar to Burstein and Vogel (2015), X_{iln} does not have closed-form expression due to technology-capital complementarity. Consumers in market n can purchase goods produced by firms from all different origins and thus the price index is

$$P_n = \left[E_\phi \left(\sum_i M_i S_{in}(\phi, a_i, b_i) \left[\frac{\sigma}{\sigma - 1} E_{\mathbf{z}} \left(\min_l C_{iln}(\phi, \mathbf{z}, a_i, b_i) \right) \right]^{1-\sigma} \right) \right]^{1/(1-\sigma)}, \tag{8}$$

where I have applied the constant markup rule under the CES demand.

A general equilibrium of the model is defined as follows.

Definition 1 (General Equilibrium). *A general equilibrium of the model is a vector of variables $\{(a_i, b_i), r_i, w_i, P_i, X_i, M_i\}$ such that*

1. *Firms choose optimal technologies to maximize global expected profit*

$$(a_i, b_i) = \arg \max_{(a,b) \in \Theta} E_\phi [\pi_i(\phi, a, b)] \tag{9}$$

2. *Net profit is non-positive due to free entry* $E_\phi[\pi_i(\phi,a,b)] - P_i F_{ei} \le 0$*, and* $E_\phi[\pi_i(\phi,a,b)] - P_i F_{ei} = 0$ *when* $M_i > 0$*.*

3. *Capital and labor markets clear:*

$$K_i = \frac{1}{\tilde{\sigma}} \sum_{j,n} M_j \int S_{jn}(\phi,a_j,b_j) X_{jin}(\phi,a_j,b_j) \frac{\kappa_{ji}(\phi,a_j,b_j)}{r_i} dF(\phi), \quad (10)$$

$$L_i = \frac{1}{\tilde{\sigma}} \sum_{j,n} M_j \int S_{jn}(\phi,a_j,b_j) X_{jin}(\phi,a_j,b_j) \frac{1-\kappa_{ji}(\phi,a_j,b_j)}{w_i} dF(\phi), \quad (11)$$

where $\kappa_{ji}(\phi,a_j,b_j)$ *is the capital share of firms producing in* i *from country* j*:*

$$\kappa_{ji}(\phi,a_j,b_j) = \left(\frac{1-\lambda}{\lambda}\phi^{\xi(\varepsilon-1)}\left(\frac{a_j}{b_j}\right)^{1-\varepsilon}\left(\frac{r_i}{w_i}\right)^{\varepsilon-1}+1\right)^{-1}. \quad (12)$$

4. *Goods market clear as follows:*

$$X_i + \Delta_i = r_i K_i + w_i L_i + P_i \sum_j M_j F_{ji} E_\phi[S_{ji}(\phi,a_i,b_i)] + M_i P_i F_{ei} \quad (13)$$

where Δ_i *is the current account surplus, which I treat as exogenous in the quantitative implementation.*

5. *The price index satisfies equation (8).*

OA.3.3 Simplification when $\xi = 0$

When there is no technology-capital complementarity, i.e., $\xi = 0$, I obtain similar expressions for aggregate variables as in ARRY. First, the normalized marginal cost ζ_{iln} no longer depends on ϕ. Neither does the profitability term Ψ_{in} and the probability of producing in l conditional on serving n, ψ_{iln}. An immediate result from equations (3) and (4) is that the productivity cutoff ϕ^*_{in} can be expressed analytically

$$\begin{aligned} \pi_{i\cdot n}(\phi) &= \frac{\tilde{\sigma}^{1-\sigma} X_n}{\sigma P_n^{1-\sigma}} \Gamma \phi^{\sigma-1} \Psi_{in}^{\frac{\sigma-1}{\chi}} \\ \phi^*_{in} &= \Gamma^{-1/(\sigma-1)} \Psi_{in}^{-1/\chi} \frac{\tilde{\sigma}}{P_n} \left(\frac{\sigma P_n F}{X_n}\right)^{1/(\sigma-1)}. \end{aligned}$$

The expected global profit of the firm can be obtained by integrating over the distribution of ϕ

$$
\begin{aligned}
E_\phi\left[\pi_i(\phi,a,b)\right] &= \sum_n \int_{\phi^*_{in}} \left(\pi_{i\cdot n}(\phi) - P_n F\right) dF_i(\phi) \\
&= \frac{(\sigma-1)\,\tilde{\sigma}^{-k}\sigma^{-\frac{k}{\sigma-1}}}{k-\sigma+1}\Gamma^{k/(\sigma-1)}\phi^k_{\min}\sum_n X_n^{\frac{k}{\sigma-1}}P_n^k\Psi_{in}(a,b)^{\frac{k}{\chi}}\left(P_nF\right)^{\frac{\sigma-k-1}{\sigma-1}}.
\end{aligned}
$$

For simplicity, I denote the expected global profit as $\pi_i(a,b)$ throughout the online appendix. Further denote the demand shifter in market n as

$$
D_n \equiv \frac{(\sigma-1)\,\tilde{\sigma}^{-k}\sigma^{-\frac{k}{\sigma-1}}}{k-\sigma+1}\Gamma^{k/(\sigma-1)}\phi^k_{\min}X_n^{\frac{k}{\sigma-1}}P_n^k\left(P_nF\right)^{\frac{\sigma-k-1}{\sigma-1}},
$$

and $\pi_i(a,b)$ can be expressed as

$$
\pi_i(a,b) = \sum_n D_n\Psi_{in}^{k/\chi}. \tag{14}
$$

To obtain gravity-type expressions for trade and MP volumes, I need to follow ARRY and assume $\phi^*_{in} \geq \phi_{\min}$, i.e., for any each destination market n (including the domestic market i), the least productive firms from i do not serve that market. Under this assumption, the total sales to market n made by country i firms producing in country l is

$$
\begin{aligned}
X_{iln} &= M_i\int S_{in}(\phi)\,X_{iln}(\phi)\,dF(\phi) \\
&= M_i\int_{\phi^*_{in}}\sigma\psi_{iln}\pi_{i\cdot n}(\phi)\,dF(\phi) \\
&= M_i\int_{\phi^*_{in}}\sigma\psi_{iln}\frac{\tilde{\sigma}^{1-\sigma}X_n}{\sigma P_n^{1-\sigma}}\Gamma\phi^{\sigma-1}\Psi_{in}^{\frac{\sigma-1}{\chi}}\,dF(\phi)
\end{aligned}
$$

where $F(\phi) = 1-(\phi/\phi_{\min})^{-k}$ therefore the density function for ϕ is $k\phi^k_{\min}\phi^{-k-1}$. Because the integrand is log-linear in ϕ, I obtain a closed-form expression

$$
\begin{aligned}
X_{iln} &= M_i\sigma\psi_{iln}\frac{\tilde{\sigma}^{1-\sigma}X_n}{\sigma P_n^{1-\sigma}}\Gamma\Psi_{in}^{\frac{\sigma-1}{\chi}}k\phi^k_{\min}\int_{\phi^*_{in}}\phi^{\sigma-k-2}d\phi \\
&= M_i\Gamma\frac{k\phi^{\sigma-1}_{\min}}{k+1-\sigma}\frac{\tilde{\sigma}^{1-\sigma}X_n}{P_n^{1-\sigma}}\psi_{iln}\Psi_{in}^{\frac{\sigma-1}{\chi}},
\end{aligned}
$$

which requires a parameter restriction $k > \sigma - 1$. This is essentially the same as the aggregate flow derived in ARRY. The only difference is that the unit cost inside the expressions ψ_{iln} and Ψ_{in} depends on two factor prices (wages and rental rates) as well as the optimal technologies (a_i, b_i).

OA.3.4 Propositions and Proofs

OA.3.4.1 Lemma on the Optimal Technology Choice

Before I prove Proposition 1, I prove the following lemma:

Lemma. *The optimal technology choice* (a^*, b^*) *must be an interior point on the boundary of the technology menu, i.e.,*

$$\theta(a^*, b^*) = 1, \quad a^* \neq 0, b^* \neq 0,$$

and

$$\frac{\partial \pi_i(a_i^*, b_i^*)/\partial a}{\partial \pi_i(a_i^*, b_i^*)/\partial b} = \frac{\partial \theta(a_i^*, b_i^*)/\partial a}{\partial \theta(a_i^*, b_i^*)/\partial b}, i = N, S.$$

Proof. Recall that the production function is strictly increasing in a and b

$$q = zF(aK, bL).$$

Thus, the marginal cost strictly decreases with a and b. So the expected global profit strictly increases in a and b. On the other hand, the function that determines the boundary of the technology menu, $\theta(a, b)$, strictly increases in a and b. Consider (a, b) such that $\theta(a, b) = \lambda < 1$, then there exists $(a', b') > (a, b)$ such that $\theta(a', b') < 1$. It allows the firm to have a lower cost of production in any country, which increases the profit. Thus, the optimal technology choice (a, b) must occur on the boundary.

To show that the optimal technology must be in the interior, I denote $\delta = (\varepsilon - 1)\log(a/b)$. Since $\theta(a, b) = 1$, one can express (a, b) using δ

$$a = e^{\delta/(\varepsilon-1)}\left[1 + \exp\left(\frac{1-\eta}{1-\varepsilon}\delta\right)\right]^{1/(\eta-1)}, \quad b = \left[1 + \exp\left(\frac{1-\eta}{\varepsilon-1}\delta\right)\right]^{1/(\eta-1)}.$$

The unit cost of producing with factor prices (r, w) can be written as

$$\begin{aligned} C(a, b) &= b^{-1}\left[(r/(a/b))^{1-\varepsilon} + w^{1-\varepsilon}\right]^{1/(1-\varepsilon)} \\ &= \left[1 + \exp\left(\frac{1-\eta}{\varepsilon-1}\delta\right)\right]^{1/(1-\eta)}\left[r^{1-\varepsilon}e^{\delta} + w^{1-\varepsilon}\right]^{1/(1-\varepsilon)}. \end{aligned}$$

Taking derivative of $\log C$ with respect to δ, we have

$$\frac{\partial \log C}{\partial \delta} = \frac{1}{1-\varepsilon} \left[\kappa\left(\delta\right) - t\left(\delta\right) \right],$$

where $\kappa(\delta)$ is the capital share in production

$$\kappa(\delta) \equiv \frac{r^{1-\varepsilon} e^{\delta}}{r^{1-\varepsilon} e^{\delta} + w^{1-\varepsilon}},$$

and $t(\delta)$ is related to the technology menu and can be expressed as

$$t(\delta) \equiv \frac{\exp\left(\frac{1-\eta}{\varepsilon-1}\delta\right)}{1+\exp\left(\frac{1-\eta}{\varepsilon-1}\delta\right)}.$$

Clearly, there is a unique zero

$$\delta^{*} = \frac{(\varepsilon-1)^{2}}{\varepsilon+\eta-2} \log\left(r/w\right).$$

Consider the second order derivative,

$$\begin{aligned} \frac{\partial^{2} \log C}{\partial \delta^{2}} &= \frac{1}{1-\varepsilon} \left(\frac{\partial \kappa\left(\delta\right)}{\partial \delta} - \frac{\partial t\left(\delta\right)}{\partial \delta} \right) \\ &= \frac{1}{1-\varepsilon} \kappa(\delta)(1-\kappa(\delta)) + \frac{1-\eta}{(1-\varepsilon)^{2}} t(\delta)(1-t(\delta)). \end{aligned}$$

Since $\kappa(\delta^{*}) = t(\delta^{*})$, we have

$$\frac{\partial^{2} \log C}{\partial \delta^{2}}|_{\delta=\delta^{*}} = -\frac{\varepsilon+\eta-2}{\left(\varepsilon-1\right)^{2}} \kappa\left(\delta^{*}\right)\left(1-\kappa\left(\delta^{*}\right)\right),$$

which has the same sign as $2-\varepsilon-\eta$. Under the restriction $\varepsilon+\eta<2$, we must have $\partial^{2} \log C/\partial \delta^{2} > 0$. Since δ^{*} is the unique zero, C increases with δ when $\delta > \delta^{*}$, and decreases with δ when $\delta < \delta^{*}$. [6]

[6] When $\eta+\varepsilon>2$, δ^{*} is the local maximum instead of the local minimum of $C(\delta)$. One can calculate that $C(-\infty)=w, C(\infty)=r$ if $(\varepsilon-1)(\eta-1)<0$, and $C(-\infty)=C(\infty)=0$ when $\varepsilon>1, \eta>1$.

Now, recall that the firm is maximizing global expected profit by choosing δ

$$\pi_i(\delta) = \sum_n D_n \left(\sum_l T_{il} (\gamma_{il} C_l(\delta) \tau_{ln})^{-\chi} \right)^{k/\chi}.$$

We denote $\delta_l \equiv \arg\min C_l(\delta)$ as the optimal technology if a firm only produces in country l. We have shown that δ_l must be interior solutions. Therefore, we can find the minimum and maximum of these values, and both are interior solutions. Denote the most and least capital intensive technology as $\bar{\delta}$ and $\underline{\delta}$, respectively. Given our result above, we can show that, when $\delta < \underline{\delta}$, C_l increases with δ for all l, and when $\delta > \bar{\delta}$, C_l decreases with δ for all l. Since π_i is monotonic in all C_l, we can conclude that the value of δ that maximizes π_i must be in the range of $[\underline{\delta}, \bar{\delta}]$, thus an interior solution.

Now we want to characterize the first order condition using (a, b). The first order conditions imply the iso-profit curve is tangent to the technology boundary at any optimal technology (a_i^*, b_i^*). Since both θ and π_i are continuously differentiable and strictly increasing in (a, b), I invert $\theta(a, b) = 1$ and $\pi_i(a, b) = \pi_i(a_i^*, b_i^*)$ and get[7]

$$b = \theta^{-1}(a), \, b = \pi_i^{-1}(a).$$

Consider $f_i(a) \equiv \pi_i^{-1}(a) - \theta^{-1}(a)$, which is a continuously differentiable function in a (because C and θ are continuously differentiable). By construction, $f_i(a_i^*) = b_i^* - b_i^* = 0$. In addition, for any value of a, we must have $f(a) \geq 0$; otherwise, $\left(a, \pi_i^{-1}(a)\right)$ is in the interior of the technology menu and we can always find a technology that delivers a higher expected global profit than $\pi_i(a_i^*, b_i^*)$. Thus, $f'(a)$ must be zero at a_i^*. This implies $\frac{\partial}{\partial a}\pi_i^{-1}(a_i^*) - \frac{\partial}{\partial a}\theta^{-1}(a_i^*) = 0$. Since $\theta(a, \theta^{-1}(a)) = 1$, $\pi_i\left(a, \pi_i^{-1}(a)\right) = \pi_i(a_i^*, b_i^*)$, I obtain at (a_i^*, b_i^*)

$$\frac{\partial \pi_i / \partial a}{\partial \pi_i / \partial b} = -\frac{\partial}{\partial a}\pi_i^{-1}(a_i^*) = -\frac{\partial}{\partial a}\theta^{-1}(a_i^*) = \frac{\partial \theta / \partial a}{\partial \theta / \partial b}.$$

□

OA.3.4.2 Proof of Proposition 1

Proof of Proposition 1. In this proof, I first show that if part (1) holds, i.e., the North has relatively cheap capital $r_N/w_N < r_S/w_S$, then part (2) and (3) of the proposition must hold. The results simply come from the firm's problem of choosing optimal technology (a, b), taking

[7] Suppose $\exists\, b_1 \neq b_2$ such that $E_\phi[\pi_i(\phi, a, b_1)] = E_\phi[\pi_i(\phi, a, b_2)]$. By mean value theorem, $\exists\, \tilde{b}$ between b_1 and b_2 such that $\partial E_\phi\left[\pi_i\left(\phi, a, \tilde{b}\right)\right] / \partial b = \left[E_\phi[\pi_i(\phi, a, b_2)] - E_\phi[\pi_i(\phi, a, b_1)]\right] / (b_2 - b_1) = 0$ which contradicts $\partial E_\phi[\pi_i(\phi, a, b)] / \partial b > 0$.

aggregate demand and factor prices as given. I then prove that the factor prices must be such that $r_N/w_N < r_S/w_S$ otherwise factor markets will not clear.

(Part 2 and 3) Under the assumption that $\xi = 0$, I can obtain a closed-form expression of the global expected profit as in equation (14). When there are N_N symmetric Northern countries and N_S symmetric Southern countries, the expected global profit of a Northern firm can be simplified as

$$\pi_N(a,b) = D_N \Psi_{Nh}^{k/\chi} + (N_N - 1) D_N \Psi_{NN}^{k/\chi} + N_S D_S \Psi_{NS}^{k/\chi}, \tag{15}$$

where

$$\begin{aligned} \Psi_{Nh} &\equiv C_N^{-\chi} + (N_N - 1)(\gamma C_N \tau)^{-\chi} + N_S (\gamma C_S \tau)^{-\chi}, \\ \Psi_{NN} &\equiv (C_N \tau)^{-\chi} + (\gamma C_N)^{-\chi} + (N_N - 2)(\gamma C_N \tau)^{-\chi} + N_S (\gamma C_S \tau)^{-\chi}, \\ \Psi_{NS} &\equiv (C_N \tau)^{-\chi} + (N_N - 1)(\gamma C_N \tau)^{-\chi} + (\gamma C_S)^{-\chi} + (N_S - 1)(\gamma C_S \tau)^{-\chi}. \end{aligned} \tag{16}$$

The first term in equation (15) is the expected profit from the domestic market, the second term is the total expected profit from other Northern markets, and the third term is the total expected profit from selling to the South. The expressions for Ψ's are also intuitive. Ψ_{Nh} determines the expected profit from serving the home market through (1) selling from itself (2) selling from a subsidiary in another Northern country and (3) selling from a subsidiary in a Southern country. The expressions for Ψ_{NN} and Ψ_{NS} can be interpreted in the same way. The global expected profit depends on technology choice through the production costs C_N and C_S. For a typical Southern firm $i \in S$, the expression for $\pi_S(a,b)$ is similar.

Taking derivatives of π_i ($i \in \{N, S\}$) with respect to C_N and C_S, and one can write

$$\frac{\partial \pi_i / \partial C_N}{\partial \pi_i / \partial C_S} = \frac{A_i}{B_i} \left(\frac{C_N}{C_S} \right)^{-\chi - 1}, \; i \in \{N, S\},$$

where A_i and B_i are functions of γ, τ, D_N, D_S, N_N, N_S and Ψ_{in}.

One can verify that, the expressions of A_N, B_N can be written as follows:

$$\begin{aligned} A_N &= D_N \left(1 + (N_N - 1)(\gamma\tau)^{-\chi}\right) \Psi_{Nh}^{k/\chi - 1} + D_S N_S \tau^{-\chi} \left(1 + (N_N - 1)\gamma^{-\chi}\right) \Psi_{NS}^{k/\chi - 1} \\ &+ D_N (N_N - 1)\left(\tau^{-\chi} + \gamma^{-\chi} + (N_N - 2)(\gamma\tau)^{-\chi}\right) \Psi_{NN}^{k/\chi - 1}, \\ B_N &= D_N N_S (\gamma\tau)^{-\chi} \Psi_{Nh}^{k/\chi - 1} + D_S N_S \gamma^{-\chi} \left(1 + (N_S - 1)\tau^{-\chi}\right) \Psi_{NS}^{k/\chi - 1} \\ &+ D_N (N_N - 1) N_S (\gamma\tau)^{-\chi} \Psi_{NN}^{k/\chi - 1}. \end{aligned}$$

By symmetry, one can obtain A_S, B_S by replacing N with S and S with N in the expressions of B_N, A_N, respectively. To see the sign of $A_N B_S - B_N A_S$, collecting terms

1. The coefficient before D_S^2 is

$$N_S\tau^{-\chi}\left(1+(N_N-1)\gamma^{-\chi}\right)\Psi_{NS}^{k/\chi-1}\times\left(\begin{array}{c}\left(1+(N_S-1)(\gamma\tau)^{-\chi}\right)\Psi_{Sh}^{k/\chi-1}\\ +(N_S-1)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_S-2)(\gamma\tau)^{-\chi}\right)\Psi_{SS}^{k/\chi-1}\end{array}\right)$$
$$-N_S\gamma^{-\chi}\left(1+(N_S-1)\tau^{-\chi}\right)\Psi_{NS}^{k/\chi-1}\times\left(N_N(\gamma\tau)^{-\chi}\Psi_{Sh}^{k/\chi-1}+(N_S-1)N_N(\gamma\tau)^{-\chi}\Psi_{SS}^{k/\chi-1}\right).$$

Further collecting terms,

(a) The coefficient before $\Psi_{Sh}^{k/\chi-1}\Psi_{NS}^{k/\chi-1}$ is

$$\left(1+(N_S-1)(\gamma\tau)^{-\chi}\right)N_S\tau^{-\chi}\left(1+(N_N-1)\gamma^{-\chi}\right)-N_N(\gamma\tau)^{-\chi}N_S\gamma^{-\chi}\left(1+(N_S-1)\tau^{-\chi}\right)$$
$$=N_S\tau^{-\chi}\left[\begin{array}{c}\left(1+(N_N-1)\gamma^{-\chi}\right)\left(1+(N_S-1)(\gamma\tau)^{-\chi}\right)\\ -N_N\gamma^{-\chi}\left(\gamma^{-\chi}+(N_S-1)(\gamma\tau)^{-\chi}\right)\end{array}\right]>0.$$

(b) the coefficient before $\Psi_{SS}^{k/\chi-1}\Psi_{NS}^{k/\chi-1}$ is

$$(N_S-1)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_S-2)(\gamma\tau)^{-\chi}\right)N_S\tau^{-\chi}\left(1+(N_N-1)\gamma^{-\chi}\right)$$
$$-(N_S-1)N_N(\gamma\tau)^{-\chi}N_S\gamma^{-\chi}\left(1+(N_S-1)\tau^{-\chi}\right)$$
$$=(N_S-1)N_S\tau^{-\chi}\left(\begin{array}{c}\left(1+(N_N-1)\gamma^{-\chi}\right)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_S-2)(\gamma\tau)^{-\chi}\right)\\ -N_N\gamma^{-\chi}\left(\gamma^{-\chi}+(N_S-1)(\gamma\tau)^{-\chi}\right)\end{array}\right)$$
$$>0$$

Thus, when $\gamma>1$, the coefficient before D_S^2 is positive.

2. Similarly, one can show that the coefficient before D_N^2 is also positive. This can also be shown using an symmetry argument - the coefficient before D_N^2 is simply replacing N with S and S with N in the coefficient before D_S^2.

3. The coefficient before $D_N D_S$ equals

$$\left[\begin{array}{c}\left(1+(N_N-1)(\gamma\tau)^{-\chi}\right)\Psi_{Nh}^{k/\chi-1}\\+(N_N-1)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_N-2)(\gamma\tau)^{-\chi}\right)\Psi_{NN}^{k/\chi-1}\end{array}\right]$$
$$\times\left[\begin{array}{c}\left(1+(N_S-1)(\gamma\tau)^{-\chi}\right)\Psi_{Sh}^{k/\chi-1}\\+(N_S-1)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_S-2)(\gamma\tau)^{-\chi}\right)\Psi_{SS}^{k/\chi-1}\end{array}\right]$$
$$+N_S\tau^{-\chi}\left(1+(N_N-1)\gamma^{-\chi}\right)\Psi_{NS}^{k/\chi-1}\times N_N\tau^{-\chi}\left(1+(N_S-1)\gamma^{-\chi}\right)\Psi_{SN}^{k/\chi-1}$$
$$-\left(N_S(\gamma\tau)^{-\chi}\Psi_{Nh}^{\frac{-\rho}{1-\rho}}+(N_N-1)N_S(\gamma\tau)^{-\chi}\Psi_{NN}^{k/\chi-1}\right)$$
$$\times\left(N_N(\gamma\tau)^{-\chi}\Psi_{Sh}^{\frac{-\rho}{1-\rho}}+(N_S-1)N_N(\gamma\tau)^{-\chi}\Psi_{SS}^{k/\chi-1}\right)$$
$$-N_S\gamma^{-\chi}\left(1+(N_S-1)\tau^{-\chi}\right)\Psi_{NS}^{k/\chi-1}\times N_N\gamma^{-\chi}\left(1+(N_N-1)\tau^{-\chi}\right)\Psi_{SN}^{k/\chi-1}.$$

Collecting terms this coefficient becomes

$$\begin{array}{l}\lambda_{Nh,Sh}\Psi_{Nh}^{k/\chi-1}\Psi_{Sh}^{k/\chi-1}+\lambda_{Nh,SS}\Psi_{Nh}^{k/\chi-1}\Psi_{SS}^{k/\chi-1}+\lambda_{NN,Sh}\Psi_{NN}^{k/\chi-1}\Psi_{Sh}^{k/\chi-1}\\+\lambda_{NN,SS}\Psi_{NN}^{k/\chi-1}\Psi_{SS}^{k/\chi-1}+\lambda_{NS,SN}\Psi_{NS}^{k/\chi-1}\Psi_{SN}^{k/\chi-1},\end{array}\tag{17}$$

where λ's are coefficients that involve γ, τ, N_N, N_S.

(a) the coefficient before $\Psi_{Nh}^{k/\chi-1}\Psi_{Sh}^{k/\chi-1}$ is

$$\begin{array}{l}\lambda_{Nh,Sh}\equiv\left(1+(N_N-1)(\gamma\tau)^{-\chi}\right)\left(1+(N_S-1)(\gamma\tau)^{-\chi}\right)-\left(N_S(\gamma\tau)^{-\chi}\right)\left(N_N(\gamma\tau)^{-\chi}\right)\\=\left(1-(\gamma\tau)^{-\chi}\right)\left(1+(N_N+N_S-1)(\gamma\tau)^{-\chi}\right)>0\end{array}$$

(b) the coefficient before $\Psi_{Nh}^{k/\chi-1}\Psi_{SS}^{k/\chi-1}$ is

$$\begin{array}{lcl}\lambda_{Nh,SS}&\equiv&\left(1+(N_N-1)(\gamma\tau)^{-\chi}\right)(N_S-1)\times\left(\tau^{-\chi}+\gamma^{-\chi}+(N_S-2)(\gamma\tau)^{-\chi}\right)\\&-&N_S(\gamma\tau)^{-\chi}(N_S-1)N_N(\gamma\tau)^{-\chi}\\&=&(N_S-1)\left(1-(\gamma\tau)^{-\chi}\right)\left(\tau^{-\chi}+\gamma^{-\chi}-2(\gamma\tau)^{-\chi}\right)\\&+&(N_S-1)\left[\left(1-(\gamma\tau)^{-\chi}\right)N_S+\left(\tau^{-\chi}+\gamma^{-\chi}-2(\gamma\tau)^{-\chi}\right)N_N\right](\gamma\tau)^{-\chi}\\&>0&\end{array}$$

note here we use the fact that

$$\tau^{-\chi}+\gamma^{-\chi}-2(\gamma\tau)^{-\chi}=-2\underbrace{\left(\gamma^{-\chi}-\frac{1}{2}\right)}_{\in(-1/2,1/2)}\underbrace{\left(\tau^{-\chi}-\frac{1}{2}\right)}_{\in(-1/2,1/2)}+\frac{1}{2}\in(0,1)$$

(c) the coefficient before $\Psi_{NN}^{k/\chi-1}\Psi_{Sh}^{k/\chi-1}$ is

$$\begin{aligned}\lambda_{NN,Sh} &\equiv (N_N-1)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_N-2)(\gamma\tau)^{-\chi}\right)\times\left(1+(N_S-1)(\gamma\tau)^{-\chi}\right)\\ &- (N_N-1)N_S(\gamma\tau)^{-\chi}N_N(\gamma\tau)^{-\chi}\\ &= (N_N-1)\left(1-(\gamma\tau)^{-\chi}\right)\left(\tau^{-\chi}+\gamma^{-\chi}-2(\gamma\tau)^{-\chi}\right)\\ &+ (N_N-1)\left[\ \left(1-(\gamma\tau)^{-\chi}\right)N_N+\left(\tau^{-\chi}+\gamma^{-\chi}-2(\gamma\tau)^{-\chi}\right)N_S\ \right](\gamma\tau)^{-\chi}\\ &>0\end{aligned}$$

(d) the coefficient before $\Psi_{NN}^{k/\chi-1}\Psi_{SS}^{k/\chi-1}$ is

$$\begin{aligned}\lambda_{NN,SS} &\equiv (N_N-1)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_N-2)(\gamma\tau)^{-\chi}\right)\times(N_S-1)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_S-2)(\gamma\tau\right.\\ &- (N_N-1)\left(N_S(\gamma\tau)^{-\chi}\right)(N_S-1)\left(N_N(\gamma\tau)^{-\chi}\right)\\ &= (N_N-1)(N_S-1)\left(\tau^{-\chi}+\gamma^{-\chi}-2(\gamma\tau)^{-\chi}\right)\times\left[\tau^{-\chi}+\gamma^{-\chi}+(N_N+N_S-2)(\gamma\tau)^{-\right.\\ &>0\end{aligned}$$

(e) the coefficient before $\Psi_{NS}^{k/\chi-1}\Psi_{SN}^{k/\chi-1}$ is

$$\begin{aligned}\lambda_{NS,SN} &\equiv N_S\tau^{-\chi}\left(1+(N_N-1)\gamma^{-\chi}\right)N_N\tau^{-\chi}\left(1+(N_S-1)\gamma^{-\chi}\right)\\ &- N_S\gamma^{-\chi}\left(1+(N_S-1)\tau^{-\chi}\right)N_N\gamma^{-\chi}\left(1+(N_N-1)\tau^{-\chi}\right)\\ &= N_SN_N\left(\tau^{-\chi}-\gamma^{-\chi}\right)\left[\tau^{-\chi}+\gamma^{-\chi}+(N_N+N_S-2)(\gamma\tau)^{-\chi}\right],\end{aligned}$$

which has the same sign as $\gamma-\tau$.

Combining all the results above, one can show that, under the assumption $\gamma>\tau\geq 1$,

$$A_NB_S-B_NA_S>0.$$

Note that this is not the only sufficient condition for $A_NB_S>B_NA_S$. For example, one can verify that when $\tau\to\infty,\gamma>1$, we also have $A_NB_S>B_NA_S$.

Therefore, under either assumptions, we have

$$\frac{\partial \pi_N / \partial C_N}{\partial \pi_N / \partial C_S} > \frac{\partial \pi_S / \partial C_N}{\partial \pi_S / \partial C_S}.$$

for any (C_N, C_S). Intuitively, firms care relatively more about the production cost in its home region. This is because the foreign MP cost γ is greater than 1, which makes the firm put more weight on the profit obtained from domestic production when choosing the optimal technology.

Since both Northern and Southern firms face the same technology menu, to compare their technology choices, one just needs to compare the relative impact of a and b on profits of Northern and Southern firms. Using the chain rule,

$$\frac{\partial \pi_N}{\partial a} = \frac{\partial \pi_N}{\partial C_N} \frac{\partial C_N}{\partial a} + \frac{\partial \pi_N}{\partial C_S} \frac{\partial C_S}{\partial a},$$

and the expressions for $\partial \pi_N / \partial b$, $\partial \pi_S / \partial a$ and $\partial \pi_S / \partial b$ are similar.

Denote $\sigma_N \equiv \frac{\partial \pi_N / \partial C_N}{\partial \pi_N / \partial C_S}$ and $\sigma_S \equiv \frac{\partial \pi_S / \partial C_N}{\partial \pi_S / \partial C_S}$, the difference between $\frac{\partial \pi_N / \partial a}{\partial \pi_N / \partial b}$ and $\frac{\partial \pi_S / \partial a}{\partial \pi_S / \partial b}$ is

$$\frac{\partial \pi_N / \partial a}{\partial \pi_N / \partial b} - \frac{\partial \pi_S / \partial a}{\partial \pi_S / \partial b} = \frac{\partial C_N}{\partial b} \frac{\partial C_S}{\partial b} \frac{(\sigma_N - \sigma_S) \left(\frac{\partial C_N / \partial a}{\partial C_N / \partial b} - \frac{\partial C_S / \partial a}{\partial C_S / \partial b} \right)}{\left(\sigma_N \frac{\partial C_N}{\partial b} + \frac{\partial C_S}{\partial b} \right) \left(\sigma_S \frac{\partial C_N}{\partial b} + \frac{\partial C_S}{\partial b} \right)}.$$

Since $\sigma_N > \sigma_S$, the sign of the above expression is determined by the second term in the numerator. Under CES production function,

$$\frac{\partial C_N / \partial a}{\partial C_N / \partial b} = \frac{\lambda}{1 - \lambda} \left(\frac{r_N}{w_N} \right)^{1-\varepsilon} \left(\frac{a}{b} \right)^{\varepsilon - 2}.$$

Since for this part, I assume $r_N / w_N < r_S / w_S$, thus $\frac{\partial C_N / \partial a}{\partial C_N / \partial b} \lesseqgtr \frac{\partial C_S / \partial a}{\partial C_S / \partial b}$ if $\varepsilon \lesseqgtr 1$.

Now I can characterize the optimal technology under the constraint $\theta(a, b) = 1$. First consider the case $\varepsilon < 1$ and denote $(a_N, b_N) = \arg\max \pi_N$, $(a_S, b_S) = \arg\max \pi_S$. Suppose $a_N \geq a_S$. Similar to the proof of the lemma, I invert the two iso-profit curves $\pi_i(a, b) = \pi_i(a_i, b_i), i = N, S$ and get $b = \pi_i^{-1}(a)$. Define a function

$$g(a) \equiv \pi_N^{-1}(a) - \pi_S^{-1}(a)$$

and its derivative

$$g'(a) = -\frac{\partial \pi_N / \partial a}{\partial \pi_N / \partial b} + \frac{\partial \pi_S / \partial a}{\partial \pi_S / \partial b} > 0,$$

which I have proved to be positive above.

On the other hand, $g(a_S) = \pi_N^{-1}(a_S) - b_S = \pi_N^{-1}(a_S) - \theta^{-1}(a_S) \geq 0$, and $g(a_N) = b_N - \pi_S^{-1}(a_N) \leq 0$. $g(a)$ is strictly increasing implies $a_S \geq a_N$. Suppose $a_S = a_N$, then both the Northern and Southern firms choose exactly the same technology $\left(\tilde{a}, \tilde{b}\right)$. By the lemma I proved earlier,

$$\frac{\partial \pi_N / \partial a}{\partial \pi_N / \partial b} = \frac{\partial \theta / \partial a}{\partial \theta / \partial b} = \frac{\partial \pi_S / \partial a}{\partial \pi_S / \partial b},$$

which contradicts $\frac{\partial \pi_N / \partial a}{\partial \pi_N / \partial b} < \frac{\partial \pi_S / \partial a}{\partial \pi_S / \partial b}$. Thus, a_S must be greater than a_N. Since the optimal technology occurs along $\theta(a, b) = 1$, one immediately gets $b_S < b_N$. The relative demand for capital and labor for firms from country i is

$$\frac{K(r_l, w_l; a_i, b_i)}{L(r_l, w_l; a_i, b_i)} = \frac{\lambda}{1-\lambda}\left(\frac{a_i}{b_i}\right)^{\varepsilon-1}\left(\frac{r_l}{w_l}\right).$$

Since $a_S > a_N$, $b_S < b_N$, the relative demand for Northern firms must be larger than those of the Southern firms when facing the same factor prices (r_l, w_l).

When $\varepsilon > 1$, one can prove $a_S < a_N$, $b_S > b_N$ and again

$$\frac{K(r_l, w_l; a_N, b_N)}{L(r_l, w_l; a_N, b_N)} > \frac{K(r_l, w_l; a_S, b_S)}{L(r_l, w_l; a_S, b_S)}.$$

This completes the proof for part (2) of the proposition assuming $r_N/w_N < r_S/w_S$ holds.

Next, I prove part (3) of the proposition – the cost of production is relatively small in the domestic market, comparing to a firm originated in the opposite region. I first prove that

$$\frac{C(r_N, w_N; a_N, b_N)}{C(r_N, w_N; a_S, b_S)} \leq \frac{C(r_S, w_S; a_N, b_N)}{C(r_S, w_S; a_S, b_S)}. \tag{18}$$

Using the cost function dual to the CES production function, I have

$$C(r_l, w_l; a_i, b_i) = \frac{w_l}{b_i}\left[\lambda\left(\frac{r_l b_i}{w_l a_i}\right)^{1-\varepsilon} + 1 - \lambda\right]^{\frac{1}{1-\varepsilon}}.$$

Therefore, the inequality (18) is equivalent to

$$\ln \tilde{C}(x_{NN}) + \ln \tilde{C}(x_{SS}) \leq \ln \tilde{C}(x_{NS}) + \ln \tilde{C}(x_{SN}),$$

where $\tilde{C}(x)$ is defined as

$$\tilde{C}(x) \equiv \left(\lambda e^{(1-\varepsilon)x} + 1 - \lambda\right)^{1/(1-\varepsilon)},$$

and x_{il} is defined as

$$x_{il} \equiv \ln\left(\frac{r_l b_i}{w_l a_i}\right), \, i, l \in \{N, S\}.$$

Note that $x_{NN} + x_{SS} = x_{NS} + x_{SN}$. Suppose $\varepsilon < 1$, the technology choices satisfy $a_N/b_N < a_S/b_S$ thus $x_{NS} > x_{NN}, x_{SS} > x_{SN}$. To establish the above inequality, one just needs to show $\ln \tilde{C}(x)$ is convex in x, which can be verified using its second order derivative. Similarly, when $\varepsilon > 1$, the technology choices satisfy $a_N/b_N > a_S/b_S$ thus $x_{SS} > x_{NS}, x_{SN} > x_{NN}$. In this case, one can show $\ln \tilde{C}(x)$ is concave thus the inequality holds.

For simplicity, I denote $C_{il} = C(r_l, w_l; a_i, b_i)$. Suppose $C_{NN} > C_{SN}$, then $C_{SS} < C_{NS}$ for inequality (18) to hold. This implies that the technology of the Northern firms is strictly worse than that of the Southern firms $(C_{NN}, C_{NS}) > (C_{SN}, C_{SS})$. The Northern firms can adopt the Southern technology to improve their profits. Thus (C_{NN}, C_{NS}) is not optimal and I obtain a contradiction. Therefore, Northern firms must have cost advantage in the North and so do Southern firms in the South. Hence, I have proved part (3) conditional on part (1) of the proposition.

(Part 1) Suppose the opposite: capital is relatively cheap in the South $r_N/w_N > r_S/w_S$. In this subsection I will derive a contradiction, i.e., the relative demand for capital is larger in the South, which is inconsistent with the factor market clearing conditions.

When $r_N/w_N > r_S/w_S$, one can simply reverse the proof for part (2) and (3) and show that the Northern firms must be more labor intensive $(a_N/b_N)^{\varepsilon-1} < (a_S/b_S)^{\varepsilon-1}$ and $C_{NN}C_{SS} < C_{NS}C_{SN}$. The relative factor demand in country l is

$$\begin{aligned}
\frac{K_l}{L_l} &= \frac{\sum_{i,n} \psi_{iln} \lambda_{in}^E X_n \kappa_{il}/r_l}{\sum_{i,n} \psi_{iln} \lambda_{in}^E X_n (1-\kappa_{il})/w_l} \\
&= \left(\frac{r_l}{w_l}\right)^{-\varepsilon} \frac{\sum_{i,n} \psi_{iln} \lambda_{in}^E X_n C_{il}^{\varepsilon-1} a_i^{\varepsilon-1}}{\sum_{i,n} \psi_{iln} \lambda_{in}^E X_n C_{il}^{\varepsilon-1} b_i^{\varepsilon-1}},
\end{aligned} \tag{19}$$

where I have applied the CES form of the cost functions. λ_{in}^E denotes the share of sales in market n by country i firms producing in all potential locations

$$\lambda_{in}^E = \frac{M_i \Psi_{in}^{k/\chi}}{\sum_i M_i \Psi_{in}^{k/\chi}}.$$

Define $K_N^e \equiv \sum_{i,n} \psi_{iNn} \lambda_{in}^E X_n C_{iN}^{\varepsilon-1} a_i^{\varepsilon-1}$ (the numerator of the second term in equation (19)) and define L_N^e, K_S^e, L_S^e similarly. I next show that $K_N^e/L_N^e \leq K_S^e/L_S^e$. This, together with the assumption that $r_N/w_N \geq r_S/w_S$, implies that $K_N/L_N \geq K_S/L_S$ which contradicts with the assumption that the North is more capital abundant and both factor markets clear.

Using the formulas for ψ_{iln} and λ_{in}^E, K_N^e can be rewritten as

$$K_N^e = \sum_{i,n} T_{il} \left(\gamma_{il} \tau_{ln}\right)^{-\chi} \frac{M_i}{\sum_i M_i \Psi_{in}^{k/\chi}} \Psi_{in}^{k/\chi-1} C_{il}^{-\chi+\varepsilon-1} a_i^{\varepsilon-1}.$$

Consider a typical Northern country, its factor demand can be decomposed into: (1) home firms production and sales to (1.1) home market (1.2) another Northern market (1.3) a Souther market; (2) production of subsidiaries headquartered in another Northern country and sales to (2.1) the host (Northern) country (2.2) the source (Northern) country (2.3) a third Northern country (2.4) a Southern country; (3) production of subsidiaries headquartered in a Southern country and sales to (3.1) the source Southern country (3.2) another Southern country (3.3) the host (Northern) country (3.4) another Northern country. Collecting terms in $K_N^e L_S^e - K_S^e L_N^e$, I can show that it is always non-negative.

In particular, one can combine terms 1.1+2.2, 1.2+2.1+2.3, 1.3+2.4, 3.1, 3.2, 3.3+3.4 and get

$$\begin{aligned}
K_N^e &= M_N D_N \left(1 + (N_N - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{Nh} C_{NN}^{\varepsilon-\chi-1} a_N^{\varepsilon-1} \\
&+ M_N D_N (N_N - 1)\left(\tau^{-\chi} + \gamma^{-\chi} + (N_N - 2)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{NN} C_{NN}^{\varepsilon-\chi-1} a_N^{\varepsilon-1} \\
&+ M_N N_S D_S \left(\tau^{-\chi} + (N_N - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{NS} C_{NN}^{\varepsilon-\chi-1} a_N^{\varepsilon-1} \\
&+ N_S M_S D_S (\gamma\tau)^{-\chi} \tilde{\Psi}_{Sh} C_{SN}^{\varepsilon-\chi-1} a_S^{\varepsilon-1} \\
&+ N_S M_S (N_S - 1) D_S (\gamma\tau)^{-\chi} \tilde{\Psi}_{SS} C_{SN}^{\varepsilon-\chi-1} a_S^{\varepsilon-1} \\
&+ N_S M_S D_N \left(\gamma^{-\chi} + (N_N - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{SN} C_{SN}^{\varepsilon-\chi-1} a_S^{\varepsilon-1},
\end{aligned}$$

where we have defined

$$\tilde{\Psi}_{in} \equiv \Psi_{in}^{k/\chi-1}$$

for notational simplicity. We immediately obtain L_N^e by replacing all the capital shares with the corresponding labor shares in the above expression. Due to symmetry, it is also straightforward to get (K_S^e, L_S^e) by replacing N with S and S with N.

Next, we collect terms in $K_N^e L_S^e - K_S^e L_N^e$.

1. Coefficient before $M_N^2 D_N^2$

$$
\begin{aligned}
0 = & \left[\left(1+\left(N_N-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{Nh}+\left(N_N-1\right)\left(\tau^{-\chi}+\gamma^{-\chi}+\left(N_N-2\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NN}\right] \\
& \times\left(N_N(\gamma\tau)^{-\chi}\tilde{\Psi}_{Nh}+N_N\left(N_N-1\right)(\gamma\tau)^{-\chi}\tilde{\Psi}_{NN}\right)\left(C_{NN}C_{NS}\right)^{\varepsilon-\chi-1}\left(a_N b_N\right)^{\varepsilon-1} \\
& -\left(N_N(\gamma\tau)^{-\chi}\tilde{\Psi}_{Nh}+N_N\left(N_N-1\right)(\gamma\tau)^{-\chi}\tilde{\Psi}_{NN}\right)\left(C_{NN}C_{NS}\right)^{\varepsilon-\chi-1}\left(a_N b_N\right)^{\varepsilon-1} \\
& \times\left[\left(1+\left(N_N-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{Nh}+\left(N_N-1\right)\left(\tau^{-\chi}+\gamma^{-\chi}+\left(N_N-2\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NN}\right]
\end{aligned}
$$

Coefficient before $M_S^2 D_S^2$ is also zero due to symmetry.

2. Coefficient before $M_S^2 D_N^2$

$$
\begin{aligned}
0 = & N_S\left(\gamma^{-\chi}+\left(N_N-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SN}C_{SN}^{\varepsilon-\chi-1}a_S^{\varepsilon-1}\times N_N\left(\tau^{-\chi}+\left(N_S-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SN}C_{SS}^{\varepsilon-\chi-1}b_S^{\varepsilon-1} \\
& -N_N\left(\tau^{-\chi}+\left(N_S-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SN}C_{SS}^{\varepsilon-\chi-1}a_S^{\varepsilon-1}\times N_S M_S D_N\left(\gamma^{-\chi}+\left(N_N-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SN}C_{SN}^{\varepsilon-\chi-1}b
\end{aligned}
$$

Similarly the coefficient before $M_N^2 D_S^2$ is zero.

3. Coefficient before $M_N M_S D_N^2$

$$
\begin{aligned}
& \left[\left(1+\left(N_N-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{Nh}+\left(N_N-1\right)\left(\tau^{-\chi}+\gamma^{-\chi}+\left(N_N-2\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NN}\right] \\
& \times N_N\left(\tau^{-\chi}+\left(N_S-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SN}C_{SS}^{\varepsilon-\chi-1}C_{NN}^{\varepsilon-\chi-1}a_N^{\varepsilon-1}b_S^{\varepsilon-1} \\
& +N_S N_N\left[(\gamma\tau)^{-\chi}\tilde{\Psi}_{Nh}+\left(N_N-1\right)(\gamma\tau)^{-\chi}\tilde{\Psi}_{NN}\right]\left(\gamma^{-\chi}+\left(N_N-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SN}C_{NS}^{\varepsilon-\chi-1}C_{SN}^{\varepsilon-\chi-1}a_S^{\varepsilon-1}b_N^{\varepsilon-1} \\
& -N_S N_N\left[(\gamma\tau)^{-\chi}\tilde{\Psi}_{Nh}+\left(N_N-1\right)(\gamma\tau)^{-\chi}\tilde{\Psi}_{NN}\right]\left(\gamma^{-\chi}+\left(N_N-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SN}C_{SN}^{\varepsilon-\chi-1}C_{NS}^{\varepsilon-\chi-1}a_N^{\varepsilon-1}b_S^{\varepsilon-1} \\
& -\left[\left(1+\left(N_N-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{Nh}+\left(N_N-1\right)\left(\tau^{-\chi}+\gamma^{-\chi}+\left(N_N-2\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NN}\right] \\
& \times N_N\left(\tau^{-\chi}+\left(N_S-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SN}C_{SS}^{\varepsilon-\chi-1}C_{NN}^{\varepsilon-\chi-1}a_S^{\varepsilon-1}b_N^{\varepsilon-1} \\
& =\left(a_N^{\varepsilon-1}b_S^{\varepsilon-1}-a_S^{\varepsilon-1}b_N^{\varepsilon-1}\right)N_N\times\Xi
\end{aligned}
$$

where

$$
\begin{aligned}
\Xi \equiv & \left[\left(1+\left(N_N-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{Nh}+\left(N_N-1\right)\left(\tau^{-\chi}+\gamma^{-\chi}+\left(N_N-2\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NN}\right] \\
& \times\left(\tau^{-\chi}+\left(N_S-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SN}C_{SS}^{\varepsilon-\chi-1}C_{NN}^{\varepsilon-\chi-1} \\
& -N_S\left[(\gamma\tau)^{-\chi}\tilde{\Psi}_{Nh}+\left(N_N-1\right)(\gamma\tau)^{-\chi}\tilde{\Psi}_{NN}\right]\left(\gamma^{-\chi}+\left(N_N-1\right)\gamma\tau^{-\chi}\right)\tilde{\Psi}_{SN}C_{SN}^{\varepsilon-\tilde{k}-1}C_{NS}^{\varepsilon-\chi-1} \\
& \geq \tilde{\Psi}_{SN}C_{SN}^{\varepsilon-\chi-1}C_{NS}^{\varepsilon-\chi-1}\times\left\{\begin{array}{l}
\left[\left(1+\left(N_N-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{Nh}+\left(N_N-1\right)\left(\tau^{-\chi}+\gamma^{-\chi}+\left(N_N-2\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NN}\right] \\
\times\left(\tau^{-\chi}+\left(N_S-1\right)(\gamma\tau)^{-\chi}\right) \\
-N_S\left[(\gamma\tau)^{-\chi}\tilde{\Psi}_{Nh}+\left(N_N-1\right)(\gamma\tau)^{-\chi}\tilde{\Psi}_{NN}\right]\left(\gamma^{-\chi}+\left(N_N-1\right)(\gamma\tau)^{-\chi}\right)
\end{array}\right. \\
& = \tilde{\Psi}_{SN}C_{SN}^{\varepsilon-\chi-1}C_{NS}^{\varepsilon-\chi-1}\times\left\{\begin{array}{l}
\tau^{-\chi}\Big[\Big(1+\left(N_N-1\right)(\gamma\tau)^{-\chi}\Big)\left(1+\left(N_S-1\right)\gamma^{-\chi}\right)-N_S\gamma^{-\chi}\Big(\gamma^{-\chi}+\left(N_N-1\right)(\gamma\tau)^{-\chi} \\
+\left(N_N-1\right)\tau^{-\chi}\tilde{\Psi}_{NN}\left(\tau^{-\chi}+\gamma^{-\chi}+\left(N_N-2\right)(\gamma\tau)^{-\chi}\right)\left(1+\left(N_S-1\right)\gamma^{-\chi}\right) \\
-\left(N_N-1\right)\tau^{-\chi}\tilde{\Psi}_{NN}N_s\gamma^{-\chi}\left(\gamma^{-\chi}+\left(N_N-1\right)(\gamma\tau)^{-\chi}\right)
\end{array}\right. \\
& \geq 0
\end{aligned}
$$

Note that we used the result $C_{NN}C_{SS} < C_{SN}C_{NS}$. Since $(a_N/b_N)^{\varepsilon-1} < (a_S/b_S)^{\varepsilon-1}$, the coefficient is negative.

4. Coefficients before $M_N M_S D_S^2$

$$
\begin{aligned}
& N_N N_S\left((\gamma\tau)^{-\chi}\tilde{\Psi}_{Sh}+\left(N_S-1\right)(\gamma\tau)^{-\chi}\tilde{\Psi}_{SS}\right)\left(\gamma^{-\chi}+\left(N_S-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NS}C_{NS}^{\varepsilon-\chi-1}C_{SN}^{\varepsilon-\chi-1}a_S^{\varepsilon-1}b_N^{\varepsilon-1} \\
& +\left(\left(1+\left(N_S-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{Sh}+\left(N_S-1\right)\left(\tau^{-\chi}+\gamma^{-\chi}+\left(N_S-2\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SS}\right) \\
& \times N_S\left(\tau^{-\chi}+\left(N_N-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NS}C_{NN}^{\varepsilon-\chi-1}C_{SS}^{\varepsilon-\chi-1}a_N^{\varepsilon-1}b_S^{\varepsilon-1} \\
& -\left(\left(1+\left(N_S-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{Sh}+\left(N_S-1\right)\left(\tau^{-\chi}+\gamma^{-\chi}+\left(N_S-2\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SS}\right) \\
& \times N_S\left(\tau^{-\chi}+\left(N_N-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NS}C_{NN}^{\varepsilon-\chi-1}C_{SS}^{\varepsilon-\chi-1}a_S^{\varepsilon-1}b_N^{\varepsilon-1} \\
& -N_N N_S\left((\gamma\tau)^{-\chi}\tilde{\Psi}_{Sh}+\left(N_S-1\right)(\gamma\tau)^{-\chi}\tilde{\Psi}_{SS}\right)\left(\gamma^{-\chi}+\left(N_S-1\right)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NS}C_{NS}^{\varepsilon-\chi-1}C_{SN}^{\varepsilon-\chi-1}a_N^{\varepsilon-1}b_S^{\varepsilon-1} \\
& =\left(a_N^{\varepsilon-1}b_S^{\varepsilon-1}-a_S^{\varepsilon-1}b_N^{\varepsilon-1}\right)N_S\times\Xi
\end{aligned}
$$

where

$$
\begin{aligned}
\Xi \equiv & \Big(\left(1 + (N_S - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{Sh} + (N_S - 1)\left(\tau^{-\chi} + \gamma^{-\chi} + (N_S - 2)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{SS} \Big) \\
& \times \left(\tau^{-\chi} + (N_N - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{NS} C_{NN}^{\varepsilon-\chi-1} C_{SS}^{\varepsilon-\chi-1} \\
& - N_N \Big((\gamma\tau)^{-\chi} \tilde{\Psi}_{Sh} + (N_S - 1)(\gamma\tau)^{-\chi} \tilde{\Psi}_{SS} \Big) \left(\gamma^{-\chi} + (N_S - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{NS} C_{NS}^{\varepsilon-\chi-1} C_{SN}^{\varepsilon-\chi-1} \\
& > \tilde{\Psi}_{NS} C_{NS}^{\varepsilon-\chi-1} C_{SN}^{\varepsilon-\chi-1} \times \left[\begin{array}{c} \left(1 + (N_S - 1)(\gamma\tau)^{-\chi}\right)\left(\tau^{-\chi} + (N_N - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{Sh} \\ -N_N (\gamma\tau)^{-\chi} \left(\gamma^{-\chi} + (N_S - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{Sh} \\ + (N_S - 1)\left(\tau^{-\chi} + \gamma^{-\chi} + (N_S - 2)(\gamma\tau)^{-\chi}\right)\left(\tau^{-\chi} + (N_N - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_S \\ -N_N (N_S - 1)(\gamma\tau)^{-\chi} \left(\gamma^{-\chi} + (N_S - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{SS} \end{array} \right. \\
& > 0
\end{aligned}
$$

Thus the coefficient is negative by similar arguments.

5. Coefficient before $M_N^2 D_N D_S$

$$
\begin{aligned}
0 = & \left[\left(1 + (N_N - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{Nh} + (N_N - 1)\left(\tau^{-\chi} + \gamma^{-\chi} + (N_N - 2)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{NN} \right] \\
& \times N_N \left(\gamma^{-\chi} + (N_S - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{NS} C_{NS}^{\varepsilon-\chi-1} C_{NN}^{\varepsilon-\chi-1} a_N^{\varepsilon-1} b_N^{\varepsilon-1} \\
& + \left[(\gamma\tau)^{-\chi} \tilde{\Psi}_{Nh} + (N_N - 1)(\gamma\tau)^{-\chi} \tilde{\Psi}_{NN} \right] \\
& \times N_N N_S \left(\tau^{-\chi} + (N_N - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{NS} C_{NN}^{\varepsilon-\chi-1} C_{NS}^{\varepsilon-\chi-1} a_N^{\varepsilon-1} b_N^{\varepsilon-1} \\
& - \left[(\gamma\tau)^{-\chi} \tilde{\Psi}_{Nh} + (N_N - 1)(\gamma\tau)^{-\chi} \tilde{\Psi}_{NN} \right] \\
& \times N_N N_S \left(\tau^{-\chi} + (N_N - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{NS} C_{NN}^{\varepsilon-\chi-1} C_{NS}^{\varepsilon-\chi-1} a_N^{\varepsilon-1} b_N^{\varepsilon-1} \\
& - \left[\left(1 + (N_N - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{Nh} + (N_N - 1)\left(\tau^{-\chi} + \gamma^{-\chi} + (N_N - 2)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{NN} \right] \\
& \times N_N \left(\gamma^{-\chi} + (N_S - 1)(\gamma\tau)^{-\chi}\right) \tilde{\Psi}_{NS} C_{NS}^{\varepsilon-\chi-1} C_{NN}^{\varepsilon-\chi-1} a_N^{\varepsilon-1} b_N^{\varepsilon-1}
\end{aligned}
$$

By symmetry, coefficient before $M_S^2 D_N D_S$ is also zero.

6. Coefficient before $M_N M_S D_N D_S$

$$
\begin{aligned}
&\left[\left(1+(N_N-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{Nh}+(N_N-1)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_N-2)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NN}\right]\\
&\times\left[\left(1+(N_S-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{Sh}+(N_S-1)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_S-2)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SS}\right]C_{SS}^{\varepsilon-\chi-1}C_{NN}^{\varepsilon-\chi-1}a_N^{\varepsilon-1}b_S\\
&+N_N N_S\left[(\gamma\tau)^{-\chi}\tilde{\Psi}_{Nh}+(N_N-1)(\gamma\tau)^{-\chi}\tilde{\Psi}_{NN}\right]\times\left[(\gamma\tau)^{-\chi}\tilde{\Psi}_{Sh}+(N_S-1)(\gamma\tau)^{-\chi}\tilde{\Psi}_{SS}\right]C_{SN}^{\varepsilon-\chi-1}C_{NS}\\
&+N_N N_S\left(\tau^{-\chi}+(N_N-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NS}\times\left(\tau^{-\chi}+(N_S-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SN}C_{SS}^{\varepsilon-\chi-1}C_{NN}^{\varepsilon-\chi-1}a_N^{\varepsilon-1}b_S^{\varepsilon-1}\\
&+N_N N_S\left(\gamma^{-\chi}+(N_N-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SN}\times\left(\gamma^{-\chi}+(N_S-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NS}C_{NS}^{\varepsilon-\chi-1}C_{SN}^{\varepsilon-\chi-1}a_S^{\varepsilon-1}b_N^{\varepsilon-1}\\
&-N_N N_S\left[(\gamma\tau)^{-\chi}\tilde{\Psi}_{Nh}+(N_N-1)(\gamma\tau)^{-\chi}\tilde{\Psi}_{NN}\right]\times\left[(\gamma\tau)^{-\chi}\tilde{\Psi}_{Sh}+(N_S-1)(\gamma\tau)^{-\chi}\tilde{\Psi}_{SS}\right]C_{SN}^{\varepsilon-\chi-1}C_{NS}\\
&-\left[\left(1+(N_S-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{Sh}+(N_S-1)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_S-2)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SS}\right]\\
&\times\left[\left(1+(N_N-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{Nh}+(N_N-1)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_N-2)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NN}\right]C_{NN}^{\varepsilon-\chi-1}C_{SS}^{\varepsilon-\chi-1}a_S\\
&-N_N N_S\left(\gamma^{-\chi}+(N_S-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NS}\times\left(\gamma^{-\chi}+(N_N-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SN}C_{SN}^{\varepsilon-\chi-1}C_{NS}^{\varepsilon-\chi-1}a_N^{\varepsilon-1}b_S^{\varepsilon-1}\\
&-N_N N_S\left(\tau^{-\chi}+(N_S-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SN}\times\left(\tau^{-\chi}+(N_N-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NS}C_{NN}^{\varepsilon-\chi-1}C_{SS}^{\varepsilon-\chi-1}a_S^{\varepsilon-1}b_N^{\varepsilon-1}
\end{aligned}
$$

collecting terms, we get

$$
\begin{aligned}
&\left[\left(1+(N_N-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{Nh}+(N_N-1)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_N-2)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NN}\right]\\
&\times\left[\left(1+(N_S-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{Sh}+(N_S-1)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_S-2)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SS}\right]\\
&\times C_{SS}^{\varepsilon-\chi-1}C_{NN}^{\varepsilon-\chi-1}\left(a_N^{\varepsilon-1}b_S^{\varepsilon-1}-a_S^{\varepsilon-1}b_N^{\varepsilon-1}\right)\\
&+N_N N_S\left[(\gamma\tau)^{-\chi}\tilde{\Psi}_{Nh}+(N_N-1)(\gamma\tau)^{-\chi}\tilde{\Psi}_{NN}\right]\times\left[(\gamma\tau)^{-\chi}\tilde{\Psi}_{Sh}+(N_S-1)(\gamma\tau)^{-\chi}\tilde{\Psi}_{SS}\right]\\
&\times C_{SN}^{\varepsilon-\chi-1}C_{NS}^{\varepsilon-\chi-1}\left(a_S^{\varepsilon-1}b_N^{\varepsilon-1}-a_N^{\varepsilon-1}b_S^{\varepsilon-1}\right)\\
&+N_N N_S\left(\tau^{-\chi}+(N_N-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NS}\times\left(\tau^{-\chi}+(N_S-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SN}C_{SS}^{\varepsilon-\chi-1}C_{NN}^{\varepsilon-\chi-1}\left(a_N^{\varepsilon-1}b_S^{\varepsilon-1}-a\right.\\
&+N_N N_S\left(\gamma^{-\chi}+(N_N-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{SN}\times\left(\gamma^{-\chi}+(N_S-1)(\gamma\tau)^{-\chi}\right)\tilde{\Psi}_{NS}C_{NS}^{\varepsilon-\chi-1}C_{SN}^{\varepsilon-\chi-1}\left(a_S^{\varepsilon-1}b_N^{\varepsilon-1}-a\right.
\end{aligned}
$$

Devide the above equation by $a_N^{\varepsilon-1}b_S^{\varepsilon-1}-a_S^{\varepsilon-1}b_N^{\varepsilon-1}$, expand and collecting terms

(a) coefficient before $\tilde{\Psi}_{Nh}\tilde{\Psi}_{Sh}$

$$
\begin{aligned}
&\left(1+(N_N-1)(\gamma\tau)^{-\chi}\right)\left(1+(N_S-1)(\gamma\tau)^{-\chi}\right)C_{SS}^{\varepsilon-\chi-1}C_{NN}^{\varepsilon-\chi-1}-N_N N_S(\gamma\tau)^{-2\chi}C_{SN}^{\varepsilon-\chi-1}C_{NS}^{\varepsilon-\chi-1}\\
&>\lambda_{Nh,Sh}C_{SN}^{\varepsilon-\chi-1}C_{NS}^{\varepsilon-\chi-1}>0
\end{aligned}
$$

where we used a corollary from Proposition 1 that the optimal technology must be such that $C_{SS}C_{NN} < C_{SN}C_{NS}$

(b) coefficient before $\tilde{\Psi}_{NN}\tilde{\Psi}_{Sh}$

$$\begin{aligned} & (N_N-1)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_N-2)(\gamma\tau)^{-\chi}\right)\left(1+(N_S-1)(\gamma\tau)^{-\chi}\right)C_{SS}^{\varepsilon-\chi-1}C_{NN}^{\varepsilon-\chi-1} \\ - \quad & (N_N-1)N_N N_S(\gamma\tau)^{-2\chi}C_{SN}^{\varepsilon-\chi-1}C_{NS}^{\varepsilon-\chi-1} \\ > \quad & \lambda_{NN,Sh}C_{SN}^{\varepsilon-\chi-1}C_{NS}^{\varepsilon-\chi-1} > 0 \end{aligned}$$

(c) coefficient before $\tilde{\Psi}_{Nh}\tilde{\Psi}_{SS}$

$$\begin{aligned} & (N_S-1)\left(1+(N_N-1)(\gamma\tau)^{-\chi}\right)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_S-2)(\gamma\tau)^{-\chi}\right)C_{SS}^{\varepsilon-\chi-1}C_{NN}^{\varepsilon-\chi-1} \\ & -(N_S-1)N_N N_S(\gamma\tau)^{-2\chi}C_{SN}^{\varepsilon-\chi-1}C_{NS}^{\varepsilon-\chi-1} \\ & > \lambda_{Nh,SS}C_{SN}^{\varepsilon-\chi-1}C_{NS}^{\varepsilon-\chi-1} > 0 \end{aligned}$$

(d) coefficient before $\tilde{\Psi}_{NN}\tilde{\Psi}_{SS}$

$$\begin{aligned} & (N_N-1)(N_S-1)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_N-2)(\gamma\tau)^{-\chi}\right)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_S-2)(\gamma\tau)^{-\chi}\right)C_{SS}^{\varepsilon-\chi-1}C_{NN}^{\varepsilon-\chi} \\ & -(N_N-1)(N_S-1)N_N N_S(\gamma\tau)^{-2\chi}C_{SN}^{\varepsilon-\chi-1}C_{NS}^{\varepsilon-\chi-1} \\ & > (N_N-1)(N_S-1)\left(\tau^{-\chi}+\gamma^{-\chi}-2(\gamma\tau)^{-\chi}\right)\left[\tau^{-\chi}+\gamma^{-\chi}+(N_N+N_S-2)(\gamma\tau)^{-\chi}\right]C_{SN}^{\varepsilon-\tilde{k}-1}C_{NS}^{\varepsilon-\chi} \\ & = \lambda_{NN,SS}C_{SN}^{\varepsilon-\chi-1}C_{NS}^{\varepsilon-\chi-1} > 0 \end{aligned}$$

(e) coefficient before $\tilde{\Psi}_{NS}\tilde{\Psi}_{SN}$

$$\begin{aligned} & N_N N_S\left(\tau^{-\chi}+(N_N-1)(\gamma\tau)^{-\chi}\right)\left(\tau^{-\chi}+(N_S-1)(\gamma\tau)^{-\chi}\right)C_{SS}^{\varepsilon-\chi-1}C_{NN}^{\varepsilon-\chi-1} \\ & -N_N N_S\left(\gamma^{-\chi}+(N_N-1)(\gamma\tau)^{-\chi}\right)\left(\gamma^{-\chi}+(N_S-1)(\gamma\tau)^{-\chi}\right)C_{NS}^{\varepsilon-\chi-1}C_{SN}^{\varepsilon-\chi-1} \\ & > N_N N_S\left(\tau^{-\chi}-\gamma^{-\chi}\right)\left(\tau^{-\chi}+\gamma^{-\chi}+(N_N+N_S-2)(\gamma\tau)^{-\chi}\right)C_{NS}^{\varepsilon-\chi-1}C_{SN}^{\varepsilon-\chi-1} \\ & = \lambda_{NS,SN}C_{SN}^{\varepsilon-\chi-1}C_{NS}^{\varepsilon-\chi-1} > 0 \end{aligned}$$

Thus, the sign of the coefficient before $M_N M_S D_N D_S$ is the opposite to

$$\lambda_{Nh,Sh}\tilde{\Psi}_{Nh}\tilde{\Psi}_{Sh}+\lambda_{NN,Sh}\tilde{\Psi}_{NN}\tilde{\Psi}_{Sh}+\lambda_{Nh,SS}\tilde{\Psi}_{Nh}\tilde{\Psi}_{SS}+\lambda_{NN,SS}\tilde{\Psi}_{NN}\tilde{\Psi}_{SS}+\lambda_{NS,SN}\tilde{\Psi}_{NS}\tilde{\Psi}_{SN}$$

Under the assumption that $\gamma > \tau \geq 1$, we have shown that the coefficient before $M_N M_S D_N D_S$ is negative.

Combining all the results above, we have shown that if capital is relatively more expensive in the North ($r_N/w_N \leq r_S/w_S$), we must have

$$\frac{K_N^e}{L_N^e} < \frac{K_S^e}{L_S^e}.$$

Therefore,

$$\frac{K_N}{L_N} = \left(\frac{r_N}{w_N}\right)^{-\varepsilon} \frac{K_N^e}{L_N^e} < \left(\frac{r_S}{w_S}\right)^{-\varepsilon} \frac{K_S^e}{L_S^e} = \frac{K_S}{L_S},$$

which contradicts the assumption that the North is more capital abundant. We complete the proof.

The algebra is very involved in this proof, but the intuition is that K_l^e/L_l^e summarizes the relative demand for capital and labor determined by production shares using different technologies and capital intensity of different production technologies. Due to the iceberg MP costs $\gamma > 1$, technologies from the same region have a relatively larger share in production. Furthermore, since Southern technologies are more capital intensive, K_S^e/L_S^e is larger than K_N^e/L_N^e. Therefore, I obtain a contradiction and part (1) of the proposition is proved. □

OA.3.4.3 Extending Proposition 1 to the Case with $\xi \neq 0$ and Alternative Timing Assumptions

In this subsection, I consider the robustness of Proposition 1 in a model with technology-capital complementarity and alternative timing assumptions. In particular, I assume that firms first observe their core productivity, then choose markets to enter, then choose their technologies and finally observe the location-specific productivities. Under this setup, we have the following results

Proposition 1'. *Under the new timing assumption and the conditions of Proposition 1 (within-region symmetry, and γ and τ satisfying (1) $\gamma \geq \tau > 1$ or (2) $\tau = \infty, \gamma > 1$), I further assume that the North has relatively cheap capital $r_N/w_N < r_S/w_S$ and the core productivity of the firm ϕ is high enough such that it sells in all markets regardless of the source country. Then, in a symmetric equilibrium, an optimal technology chosen by a Northern firm with $(a_N(\phi), b_N(\phi))$ must be more capital-intensive than one chosen by a Southern firm $(a_S(\phi), b_S(\phi))$ who serves the same set of markets.*

Proof. To study the technology choice problem of the firm, the key is the global expected profit, now as a function of (a_i, b_i, ϕ), where a_i and b_i are also functions of ϕ. The problem of the last stage (choice of production locations) is not affected by the change of the timing assumptions. Therefore, the expected profit of selling in market n is the same as equation

(3). Since I focus on firms that sell in all markets, the profit of a Northern firm is

$$\pi_N(a,b,\phi) = D_N(\phi)\Psi_{Nh}(a,b,\phi)^{\frac{\sigma-1}{\chi}} + (N_N-1)\, D_N(\phi)\Psi_{NN}(a,b,\phi)^{\frac{\sigma-1}{\chi}} + N_S D_S(\phi)\Psi_{NS}(a,b,\phi)^{\frac{\sigma-1}{\chi}}, \tag{20}$$

where $D_n(\phi)$ summarises the market access of country n for firms with core productivity ϕ

$$D_n(\phi) \equiv \frac{\tilde{\sigma}^{1-\sigma} X_n}{\sigma P_n^{1-\sigma}} \Gamma\left(\frac{\chi-\sigma+1}{\chi}\right)\phi^{\sigma-1}. \tag{21}$$

The above expression is very similar to equation (15). The only differences are: (1) the market access also depends on ϕ and (2) the profitability terms Ψ_{in} depend on not only the technology choice (a, b) but also the core productivity ϕ because of technology-capital complementarity (3) the power on Ψ_{in} is now $(\sigma-1)/\chi$ rather than k/χ because I no longer need to integrate over the distribution of ϕ to obtain the expected profit. However, all the arguments in the proof of parts 2 and 3 of Proposition 1 still holds.

□

OA.4 Calibration and Counterfactuals

OA.4.1 Robustness Checks for Estimating ε

In this subsection, I provided additional robustness checks for estimating the intensive elasticity ϵ. In the paper, I choose to estimate the intensive elasticity using the following regression

$$\log\left(\frac{r_l K_a}{w_a L_a}\right) = (1-\varepsilon)\log\left(\frac{r_l}{w_l}\right) + \delta_{f\times s} + u_a, \tag{22}$$

where a denotes the affiliate, f denotes the parent firm and l and s denote the location and sector of the affiliate, respectively. This specification has the advantage that I can take care of the differences in worker quality across affiliates using their wage bills. However, I need to multiple r_l on the left hand side, which also appears on the right hand side. Measurement errors in r_l can cause mechanical correlations. An alternative strategy is to use the following specification

$$\log\left(\frac{K_a}{L_a}\right) = -\varepsilon\log\left(\frac{r_l}{w_l}\right) + \delta_{f\times s} + u_a. \tag{23}$$

I report the regressions in Table OA.20, where I also instrument the relative factor prices using endowment. The estimated elasticities are slightly larger than those estimated using equation (22), but all of them are well below one.

Table OA.20: Estimate the Intensive Elasticity ε using K_a/L_a

	(1)	(2)	(3)	(4)
$log(r_l/w_l)$	-0.592		-0.620	-0.558
	(0.115)		(0.113)	(0.118)
$log(r_l/w_a)$		-0.548		
		(0.0975)		
N	23517	23517	23517	23517
Implied ε	0.592	0.548	0.620	0.558
Assumed Affiliate Age	10	10	5	20
# of firm-industry	6423	6423	6423	6423
# of host countries	21	21	21	21
# of home countries	22	22	22	22
First-stage F	145.5	74.2	145.5	145.5

Dependent variable is log of affiliates' deflated total assets divided by number of workers. The affiliate is indicated by a, and l indicates the host country. In all specifications, relative factor prices are instrumented with host country endowment, $\log(K_l/L_l)$, and parent $\times$ NACE 4-digit industry fixed effects are controlled. Columns also differ in the affiliate ages assumed when calculating the asset deflator of each country, which are displayed at the bottom of the panel. Standard errors are two-way clustered at home and host country level.

Table OA.21: Robustness checks for the estimation of ε

Sample	North	South	All Source Countries			
	(1)	(2)	(3)	(4)	(5)	(6)
$log(r_l/w_l)$	0.479	0.512	0.519		0.492	0.554
	(0.112)	(0.105)	(0.113)		(0.110)	(0.117)
$log(r_l/w_a)$				0.498		
				(0.0995)		
log(revenue)			0.0574	0.0890	0.0582	0.0565
			(0.0119)	(0.0124)	(0.0119)	(0.0120)
N	16142	7375	22532	22532	22532	22532
Implied ε	0.521	0.488	0.481	0.502	0.508	0.446
Assumed Affiliate Age	10	10	10	10	5	20
# of firm-industry	4392	2031	6111	6111	6111	6111
# of host countries	21	21	21	21	21	21
# of home countries	11	11	22	22	22	22
First-stage F	138.5	201.4	145.0	87.5	145.0	145.0

Dependent variable is log of affiliates' capital expenditure divided by wage bills. The affiliate is indicated by a, and l indicates the host country. In all specifications, relative factor prices are instrumented with host country endowment, $\log(K_l/L_l)$, and parent $\times$ NACE 4-digit industry fixed effects are controlled. Columns also differ in the firm ages assumed when calculating the asset deflator of each country, which are displayed at the bottom of the panel. Standard errors are two-way clustered at home and host country level. "North" and "South" indicate subsamples of affiliates whose home countries' capital abundance is above/below the median.

OA.4.2 Normalization of Model Parameters

In this subsection, I show that under a special case of my model (no technology-capital complementarity, i.e., $\xi = 0$), the normalization of certain parameters of the model delivers an observationally equivalent equilibrium after the model is calibrated to the same set of moments. I prove this in a more general case of the model in which the marketing costs F are origin-destination-specific (thus denoted by F_{in}) and the lower bound of the core productivity can also be origin-specific (thus denoted by $\phi_{i,min}$).

Proposition. *Under the assumption that marketing and entry costs are paid in goods rather than capital or labor, a change from F_{ei}, F_{in}, $\phi_{i,\min}$, to F'_{ei}, F'_{in}, $\phi'_{i,\min}$ delivers an observationally equivalent equilibrium after the model is calibrated to the same set of moments if*

1. *New trade and MP costs satisfy $\gamma'_{il} = S_i^\gamma D_l^\gamma \gamma_{il}, \tau'_{ln} = \tau_{ln} D_n^\tau / D_l^\gamma$.*
2. *Changes in marketing costs satisfy $\hat{F}_{in} \equiv F'_{in}/F_{in} = S_i^F D_n^F$*
3. *Changes in model primitives must satisfy*

$$\hat{F}_{ei}^{-1} \left(D_i^\tau\right)^{\frac{(\sigma-1)k}{(2-\sigma)k-(\sigma-1)}} \left(D_i^F\right)^{\frac{k-\sigma+1}{(2-\sigma)k-(\sigma-1)}} \hat{\phi}_{i,min}^k (S_i^\gamma)^{-k} (S_i^F)^{\frac{\sigma-k-1}{\sigma-1}} = 1$$

Proof. I prove the above results using a guess-and-verify strategy. I conjecture that the new equlibrium is $(r_i, w_i, P'_i, X_i, M'_i)$ and the change in the price index and mass of entrants satisfy

$$\hat{M}_i = \hat{F}_{ei}^{-1} \left(D_i^\tau\right)^{\frac{(\sigma-1)k}{(2-\sigma)k-(\sigma-1)}} \left(D_i^F\right)^{\frac{k-\sigma+1}{(2-\sigma)k-(\sigma-1)}}, \quad (24)$$

$$\hat{P}_i = \left(D_i^\tau\right)^{-\frac{(\sigma-1)k}{(2-\sigma)k-(\sigma-1)}} \left(D_i^F\right)^{-\frac{k-\sigma+1}{(2-\sigma)k-(\sigma-1)}}. \quad (25)$$

Therefore, we have $\hat{M}_i \hat{P}_i \hat{F}_{ei} = 1$ and can rewrite condition 3 as

$$\hat{M}_i \hat{\phi}_{i,min}^k (S_i^\gamma)^{-k} (S_i^F)^{\frac{\sigma-k-1}{\sigma-1}} = 1. \quad (26)$$

Since I conjecture the factor prices (r_i, w_i), unit costs of combining capital and labor do not change: $C'_{il} = C_{il}$. Some key equilibrium variables become

$$\Psi'_{in} = \sum_l \left(\gamma'_{il} C'_{il} \tau'_{ln}\right)^{-\chi} = \left(S_i^\gamma D_n^\tau\right)^{-\chi} \Psi_{in},$$

$$\psi'_{iln} = \frac{\left(\gamma'_{il} C'_{il} \tau'_{ln}\right)^{-\chi}}{\sum_l \left(\gamma'_{il} C'_{il} \tau'_{ln}\right)^{-\chi}} = \psi_{iln}.$$

Three-way sales

$$
\begin{aligned}
X'_{iln} &= \psi'_{iln}\lambda^{E\prime}_{in}X_n \\
&= \psi_{iln}\frac{M'_i\left(\phi'_{i,\min}\right)^k\left(S^{F\prime}_i\right)^{\frac{\sigma-k-1}{\sigma-1}}\left(S^{\gamma}_i\right)^{-k}\left(\Psi_{in}\right)^{k/\chi}}{\sum_i M'_i\left(\phi'_{i,\min}\right)^k\left(S^{F\prime}_i\right)^{\frac{\sigma-k-1}{\sigma-1}}\left(S^{\gamma}_i\right)^{-k}\left(\Psi_{in}\right)^{k/\chi}}X_n
\end{aligned}
$$

where

$$
\lambda^{E\prime}_{in} = \frac{M'_i\left(\phi'_{i,\min}\right)^k\left(\Psi'_{in}\right)^{k/\chi}\left(F'_{in}\right)^{\frac{\sigma-k-1}{\sigma-1}}}{\sum_i M'_i\left(\phi'_{i,\min}\right)^k\left(\Psi'_{in}\right)^{k/\chi}\left(F'_{in}\right)^{\frac{\sigma-k-1}{\sigma-1}}}.
$$

Since

$$
\begin{aligned}
\frac{M'_i\left(\phi'_{i,\min}\right)^k\left(\Psi'_{in}\right)^{k/\chi}\left(F'_{in}\right)^{\frac{\sigma-k-1}{\sigma-1}}}{M_i\phi^k_{i,\min}\Psi^{k/\chi}_{in}F^{\frac{\sigma-k-1}{\sigma-1}}_{in}} &= \hat{M}_i\hat{\phi}^k_{i,\min}\hat{\Psi}^{k/\chi}_{in}\hat{F}^{\frac{\sigma-k-1}{\sigma-1}}_{in} \\
&= \hat{M}_i\hat{\phi}^k_{i,\min}\left(S^{\gamma}_i D^{\tau}_n\right)^{-k}\hat{F}^{\frac{\sigma-k-1}{\sigma-1}}_{in},
\end{aligned}
$$

we know that if $\hat{F}_{in} = S^F_i D^F_n$, the value of D^F_n will not affect $\lambda^{E\prime}_{in}$. However, changes in the other terms that are specific to the home country i will have an impact on $\lambda^{E\prime}_{in}$ thus on X'_{iln}. To ensure $X'_{iln} = X_{iln}$, what we need is exactly equation (26). Under this condition, X_{iln} and λ^E_{in} do not change. Neither do the trade and MP shares.

We now verify the change in the price index must satisfy equation (25) using the definition of the price index. When the the least productive firm does not enter any market (i.e. $\phi^{*\prime}_{in}, \phi^*_{in} > \phi_{i,min}$ for any i, n), the new price index is

$$
\begin{aligned}
P'_n &= \left(\frac{k}{k-\sigma+1}\Gamma^{\frac{k}{\sigma-1}}\tilde{\sigma}^{-k}\sigma^{-\frac{k-\sigma+1}{\sigma-1}}X_n^{\frac{k-\sigma+1}{\sigma-1}}\sum_i M'_i\left(\phi'_{i,\min}\right)^k\left(\Psi'_{in}\right)^{k/\chi}\left(F'_{in}\right)^{-\frac{k-\sigma+1}{\sigma-1}}\right)^{\frac{\sigma-1}{(2-\sigma)k-(\sigma-1)}} \\
&= \left(\left(D^{\tau}_n\right)^{-k}\left(D^F_n\right)^{-\frac{k-\sigma+1}{\sigma-1}}\right)^{\frac{\sigma-1}{(2-\sigma)k-(\sigma-1)}}P_n, \qquad (27)
\end{aligned}
$$

since

$$
\hat{M}_i\hat{\phi}^k_{i,\min}\hat{\Psi}^{k/\chi}_{in}\hat{F}^{-\frac{k-\sigma+1}{\sigma-1}}_{in} = \hat{M}_i\hat{\phi}^k_{i,\min}\left(S^{\gamma}_i D^{\tau}_n\right)^{-k}\left(S^F_i D^F_n\right)^{\frac{\sigma-k-1}{\sigma-1}} = 1,
$$

which holds according to equation (26)

Given that the three way sales are the same, it is straightforward to show that the capital and labor markets clear

$$K_i = \frac{1}{\tilde{\sigma}} \sum_{j,n} X_{jin} \frac{\kappa_{ji}}{r_i},$$

$$L_i = \frac{1}{\tilde{\sigma}} \sum_{j,n} X_{jin} \frac{1-\kappa_{ji}}{w_i}.$$

The key question is that whether the goods market clearing condition still holds. Recall that the goods market clearing condition was

$$X_i + \Delta_i = \frac{1}{\tilde{\sigma}} \sum_{j,n} X_{jin} + \sum_j \frac{k-\sigma+1}{k\sigma} \lambda_{ji}^E X_i + \frac{1}{\tilde{\sigma}k} \sum_n \lambda_{in}^E X_n,$$

where the first term on the RHS is final goods consumed by households, the second term is final goods used in marketing activities while the third term is final goods used in entry activities. The first term does not change since households income does not. For the third term, the total entry costs under the new equilibrium is

$$M_i' P_i' F_{ei}' = M_i P_i F_{ei} = \frac{1}{\tilde{\sigma}k} \sum_n \lambda_{in}^E X_n,$$

where we have used $\hat{M}_i \hat{P}_i \hat{F}_{ei} = 1$. This also implies that our guess of $\hat{M}_i$ and $\hat{P}_i$ are consistent with the free-entry condition.

For the second term, we first solve for the change in the cutoff core productivity $\hat{\phi}_{in}^*$:

$$\hat{\phi}_{in}^* = \hat{\Psi}_{in}^{-1/\chi} \frac{1}{\hat{P}_n} \left(\hat{P}_n \hat{F}_{in} \right)^{1/(\sigma-1)}.$$

Consider changes in the total marketing costs by country i firms for a particular destination n

$$\begin{aligned}
\hat{M}_i \hat{P}_n \hat{F}_{in} \widehat{\bar{F}_i(\phi_{in}^*)} &= \hat{M}_i \hat{\phi}_{i,min}^k \hat{P}_n \hat{F}_{in} (\hat{\phi}_{in}^*)^{-k} \\
&= \hat{M}_i \hat{\phi}_{i,min}^k (S_i^\gamma D_n^\tau)^{-k} \hat{P}_n^{\frac{(\sigma-2)k+\sigma-1}{\sigma-1}} (S_i^F D_n^F)^{\frac{\sigma-k-1}{\sigma-1}} \\
&= 1,
\end{aligned}$$

where I have applied equations (25) and (26).

To sum up, one can normalize the values of F_{ei}, $\phi_{i,min}$ and normalize F_{in} up to an original and a destination fixed effect. After a proper adjustment of the trade and MP costs, the equilibrium solved under the new set of parameters is observationally the same in that variables such as factor prices, aggregate output and consumptions, three-way sales, trade and MP shares are exactly the same. Unobserved variables such as the price index and mass of entrants will change according to equations (25) and (26). □

Now consider the case in which I normalize all marketing costs F_{in} and $\phi_{i,\min}$ to 1 (therefore, $S_i^F = D_n^F = 1$). Also, I normalize γ_{ii}, τ_{ii} and $\bar{F}_i(\phi_{ii}^*)$, but calibrate F_{ei}. We can calculate the change in model parameters if we change the targeted $\bar{F}_i(\phi_{ii}^*)$ to $\bar{F}_i(\phi_{ii}^*)'$. To see this, consider

$$\widehat{\bar{F}(\phi_{ii}^*)} = \left(\hat{\phi}_{ii}^*\right)^{-k} = \left(S_i^\gamma\right)^{-k}\left(D_i^\tau\right)^{\frac{(\sigma-1)k}{(2-\sigma)k-(\sigma-1)}} = \hat{F}_{ei},$$

where I have applied condition 3 in the proposition. Therefore, it is clear that we need to change F_{ei} by the same percentage as $\bar{F}(\phi_{ii}^*)$.

Note that matching the trade and MP shares will induce adjustment in the calibrated trade and MP costs when we change F_{ei}. One can calculate such changes using the normalization that

$$\hat{\gamma}_{ii} = S_i^\gamma D_i^\gamma = 1, \quad \hat{\tau}_{ii} = \frac{D_i^\tau}{D_i^\gamma} = 1.$$

Substituting back into the expression of $\widehat{\bar{F}(\phi_{ii}^*)}$ we have

$$S_i^\gamma = \left(\hat{F}_{ei}\right)^{-\frac{(2-\sigma)k-(\sigma-1)}{(2-\sigma)k^2}}, \quad D_i^\tau = D_i^\gamma = \left(S_i^\gamma\right)^{-1}.$$

OA.4.3 Detailed Algorithm

In this section, I discuss the adjustment algorithm used in the calibration. Intuitively, the algorithm adjusts the parameter that matters the most for each targeted moment/equilibrium condition in each iteration, and the adjustment is proportional to the differences between the predicted and targeted moments, or to the excess equilibrium conditions. Compared to conventional gradient-based method, the adjustment approach could be much faster when the system involves many unknowns. For recent applications, see Burstein and Vogel (2017) and Ravikumar et al. (2017).

In the inner loop, given a set of outer loop parameters $(\eta, \xi, k, \chi, \lambda)$, I iterate over guesses of $(\delta_i, r_i, w_i, P_i, X_i, M_i, F_{ei})$ and the trade and MP costs $\{\tau_{ln}, \gamma_{il}\}$ such that (1) all equilibrium conditions are satisfied and (2) trade and MP shares are exactly the same as those in the data and (3) the probability of a firm serving its domestic country is 0.7 in all countries (normalization). In particular, for the m$^{\text{th}}$ guess of solutions $(\delta_i^m, r_i^m, w_i^m, P_i^m, X_i^m, M_i^m, F_{ei}^m)$, I calculate the equilibrium conditions and obtain the following deviations

$$
\begin{aligned}
f_i^{foc} &= \frac{\partial}{\partial \delta} E_\phi \left[\pi_i \left(\phi, a\left(\delta_i^m\right), b\left(\delta_i^m\right) \right) \right] / F_{ei}^m P_i^m, \\
f_i^{K/L} &= \frac{K_i^{tot}}{L_i^{tot}} / \frac{K_i}{L_i} - 1, \\
f_i^L &= L_i^{tot} / L_i - 1, \\
f_i^P &= \tilde{P}_i^m / P_i^m - 1, \\
f_i^X &= \tilde{X}_i^m / X_i^m - 1, \\
f_i^\pi &= \frac{E_\phi \left(\pi_i (\phi) \right)}{F_{ei}^m P_i^m} - 1, \\
f_i^{\Pr} &= \Pr(\phi \geq \phi_{in}^*)/0.7 - 1, \\
f_{il}^M &= \lambda_{il}^M / \lambda_{il}^{M,data} - 1, \\
f_{ln}^T &= \lambda_{ln}^T / \lambda_{ln}^{T,data} - 1,
\end{aligned}
$$

where $\tilde{P}_i^m$ is the ideal price index, and $\tilde{X}_i^m$ is the total absorption implied by the m$^{\text{th}}$ guess:

$$
\begin{aligned}
\tilde{P}_i^m &= \left[E_\phi \left(\sum_j M_j^m S_{ji} \left(\phi, a(\delta_i^m), b(\delta_i^m) \right) \left[\frac{\sigma}{\sigma - 1} E_{\mathbf{z}} \left(\min_l C_{jli} \left(\phi, \mathbf{z}, a(\delta_i^m), b(\delta_i^m) \right) \right) \right]^{1-\sigma} \right) \right]^{1/(1-\sigma)}, \\
\tilde{X}_i^m &= r_i^m K_i + w_i^m L_i + P_i^m \sum_j M_j^m F E_\phi \left[S_{ji}^m \left(\phi, a(\delta_i^m), b(\delta_i^m) \right) \right] + M_i^m P_i^m F_{ei}^m - \Delta_i.
\end{aligned}
$$

The guesses are updated as follows

$$
\begin{aligned}
\delta_i^{m+1} &= \delta_i^m \times (1 + hf_i^{foc}), \\
r_i^{m+1}/w_i^{m+1} &= r_i^m/w_i^m \times (1 + hf_i^{K/L}), \\
w_i^{m+1} &= w_i^m \times (1 + hf_i^L), \\
r_i^{m+1} &= r_i^{m+1}/w_i^{m+1} \times w_i^{m+1} \\
P_i^{m+1} &= P_i^m \times (1 + hf_i^P), \\
X_i^{m+1} &= X_i^m \times (1 + hf_i^X), \\
M_i^{m+1} &= \max\left(0, M_i^m \times (1 + hf_i^\pi)\right), \\
F_{ei}^{m+1} &= F_{ei}^m \times (1 - hf_i^{\mathrm{Pr}}), \\
\gamma_{il}^{m+1} &= \gamma_{il}^m \times (1 + hf_{il}^M), \\
\tau_{ln}^{m+1} &= \tau_{ln}^m \times (1 + hf_{ln}^T),
\end{aligned}
$$

Intuitively, I increase δ if the derivative of profit with respect to δ is positive, increase prices if there are excess demands, increase trade and MP costs if the corresponding trade and MP shares are higher than those in the data, and reduce F_{ei} if the probability of a firm serving its domestic country is higher than 0.7. I set the constant h to 0.1 in the calibration.

The outer loop iterates over guesses of $(\eta, \xi, k, \chi, \lambda)$ until the five corresponding moments are matched exactly. For k and χ, I target the restricted and unrestricted trade elasticities. Note that in the inner loop, I obtain τ_{ln} as well as three-way and two-way sales $X_{iln}, X_{\cdot ln}$. Thus, I can run the following regressions to estimate the restricted and unrestricted trade elasticities

$$
\begin{aligned}
\log X_{iln} &= \delta_{il} + \delta_{in} - \beta^r \tau_{ln} + \epsilon_{iln}, \\
\log X_{\cdot ln} &= \delta_l + \delta_n - \beta^u \tau_{ln} + \epsilon_{ln}.
\end{aligned}
$$

In each iteration of the outer loop, I increase k if β^u is smaller than the targeted value and increase χ if β^r is smaller than the targeted value. I can also calculate the average labor share across countries. I increase λ if the average labor share is smaller than that in the data.

The other two parameters η and ξ are calibrated by targeting the technology origin and size effects estimated from firm-level data. After solving the inner loop, I simulate 20,000 entrants in each country. Firms choose which markets to serve and from which country to serve a particular market according to the model. Thus, for each firm, I can determine the size of a particular affiliate R_a in country l. I then select the multinational subsample and run the following regression

$$\log\left(\frac{K_a}{w_l L_a}\right) = \delta_l + \beta_1 \log\left(\frac{K_i}{L_i}\right) + \beta_2 \log(R_a) + \epsilon_a,$$

In each iteration of the outer loop, I increase η if β_1 is smaller than the targeted value and increase ξ if β_2 is smaller than the targeted value.

OA.4.4 Gravity in the Calibrated Baseline Trade and MP Costs

The calibration produces thousands of iceberg trade and MP costs, τ_{ln} and γ_{il}. In Table OA.22 I project the calibrated costs on standard gravity variables, controlling for origin and destination fixed effects. In general, these calibrated costs are positively correlated with distance, and the impact of distance is similar for trade and MP costs. Trade and MP costs are lower for countries that share borders and a common official language and for those that have former colonial relationships. These results confirm that the model can be used to back out meaningful bilateral trade and MP costs even though it is more complicated than standard models with analytical gravity equations.

Table OA.22: Gravity in τ and γ

	(1) $log(\tau_{ln})$	(2) $log(\tau_{ln})$	(3) $log(\gamma_{il})$	(4) $log(\gamma_{il})$
log(distance)	0.280	0.253	0.268	0.236
	(0.0153)	(0.0156)	(0.0150)	(0.0144)
contiguity		-0.0839		-0.0766
		(0.0268)		(0.0310)
common language		-0.0711		-0.0922
		(0.0357)		(0.0376)
colony		-0.0825		-0.141
		(0.0306)		(0.0390)
N	1332	1332	1052	1052
R^2	0.989	0.989	0.935	0.939
mean of Y	1.544		1.512	
sd of Y	1.433		0.704	

Dependent variables are either iceberg trade costs τ or MP costs γ. All regressions control host and home country fixed effects. Standard errors clustered at host-country level. γ_{il} is set at infinity for country pairs with zero MP, which are automatically excluded from the regressions.

OA.4.5 MP Patterns using Differences in Income Per Capita

In Table OA.23, I replicate Table 5 in the paper by replacing the difference between home and host country's capital abundance with the difference in income levels. Comparing the first coefficient in Columns 2 and 1, one can conclude that the technology origin effect explains around 68% of the negative relationship between the bilateral MP shares and the difference in income levels.

Table OA.23: MP Shares, MP Costs and Country Income Differences

	(1) $log(\lambda_{il}^{M,base})$	(2) $log(\lambda_{il}^{M,TCC})$	(3) $log(\gamma_{il}^{base})$	(4) $log(\lambda_{il}^{M,base})$	(5) $log(\lambda_{il}^{M,TCC})$	(6) $log(\gamma_{il}^{base})$
diff in income	-0.765	-0.248	0.115	-0.807	-0.291	0.121
	(0.151)	(0.150)	(0.0291)	(0.149)	(0.145)	(0.0282)
log(dist)	-1.698	-1.733	0.365	-1.811	-1.856	0.382
	(0.0983)	(0.103)	(0.0184)	(0.116)	(0.122)	(0.0213)
contiguity				-1.109	-1.202	0.175
				(0.240)	(0.241)	(0.0367)
common language				0.111	0.182	-0.0336
				(0.236)	(0.230)	(0.0427)
colony				1.055	1.014	-0.161
				(0.331)	(0.326)	(0.0592)
N	1089	1089	1089	1089	1089	1089
R^2	0.784	0.781	0.904	0.793	0.790	0.907
T-stat		2.430			2.479	

Dependent variables are real/counterfactual MP shares or calibrated MP costs. $\lambda_{il}^{M,base}$ is the MP share from home country i in host country l in the data and in the baseline calibration). $\lambda^{M,TCC}$ is the counterfactual MP share when I assume all firms adopt the world average technology. γ_{il}^{base} refers to the calibrated MP costs. All regressions control host and home country fixed effects. Standard errors clustered at host-country level. Differences in country characteristics are absolute differences in log values. The T-stat is calculated based on a T-test for whether the coefficients before country differences in Columns 1 and 2 (or Columns 4 and 5) are the same. It assumes that the observations in the two regressions are independent.

OA.4.6 More Analysis of the Main Counterfactual in Section 5.1

Due to limited space, I do not report the change in real rental rates and wages from the main counterfactual in the paper. I report these results in Columns 7 and 8 of Table OA.24.

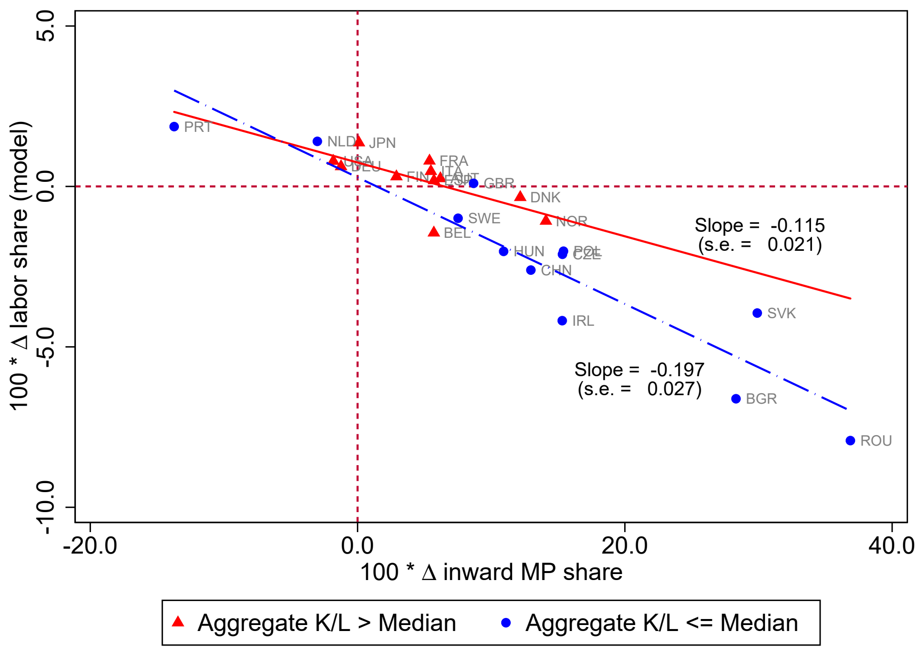
To further examine the heterogeneity in the predicted change in labor shares (Column 6 of Table OA.24), Figure OA.1 plots the predicted change in labor shares against the change in inward MP shares. I divide the sample countries into two groups: countries whose capital abundance is below/above the median. I use blue dots to represent the less capital-abundant countries and red triangles to represent the more capital-abundant ones. I also plot the best linear fit for the two groups. The estimated effect of the change in inward MP shares on the change in labor share is -0.115 (s.e. 0.021) for the more capital-abundant countries, and the effect becomes larger for the less capital-abundant countries (slope = -0.197, s.e.=0.027).

Table OA.24: Counterfactual of Reduction in Bilateral MP Costs: Additional Outcomes

ISO3	(1) Δinward MP share	(2) $\overline{\log(\hat{\gamma}_{il})}$ for each l	(3) Δoutward MP share	(4) $\overline{\log(\hat{\gamma}_{il})}$ for each i	(5) Δlabor share (model)	(6) Δlabor share (data)	(7) $\Delta\log(r/P)$	(8) $\Delta\log(w/P)$	(9) Δlog real income
AUT	6.2	-25.2		-20.5	0.2	-2.5	2.0	3.0	2.6
BEL	5.7	17.4^{i}		-6.9	-1.4	-2.1	2.0	-3.9	-1.2
BGR	28.3	-38.8		-29.6^{i}	-6.6	-7.5	13.3	-14.1	2.0
CHN	13.0	-4.2^{i}		-22.1	-2.6	-4.6	4.1	-6.8	-0.3
CZE	15.3	-22.8		-18.1	-2.1	-0.8	4.7	-3.8	0.6
DEU	-1.2	-11.6	3.7	-10.3	0.6	-5.1	0.7	3.2	2.1
DNK	12.2	-28.4		-7.5	-0.3	2.1	3.3	1.9	2.5
ESP	5.7	-15.1	6.2	-2.6	0.2	-0.3	1.0	1.7	1.4
FIN	2.9	-16.1	10.5	-12.0	0.3	3.3	2.5	3.8	3.3
FRA	5.4	-21.9	15.1	-16.1	0.8	1.2	1.0	4.3	3.0
GBR	8.7	-36.0	6.7	-20.4	0.1	0.0	1.0	1.4	1.2
HUN	10.9	-16.2	7.3	-27.5^{i}	-2.0	-4.6	8.4	0.1	4.8
IRL	15.3	-23.3		-12.4	-4.2	5.8	15.8	-1.5	8.7
ITA	5.5	-12.6	5.1	-15.1	0.5	3.1	-0.5	1.5	0.7
JPN	0.1	-14.8^{i}	7.3	-22.4	1.4	-2.1	-0.3	5.3	3.0
NLD	-3.0	-15.3		-8.0	1.4	-1.3	-8.3	-2.6	-5.3
NOR	14.1	-26.3	4.6	-1.2	-1.1	-6.4	3.3	-1.1	0.8
POL	15.4	-28.2		-22.7	-2.0	-13.3	4.3	-3.8	0.5
PRT	-13.7	-10.4	4.6	-6.0	1.9	1.2	-2.0	5.4	1.9
ROU	36.9	-34.3		-1.9^{i}	-7.9	-4.1	13.5	-19.5	0.2
SVK	29.9	-15.5		-19.2^{i}	-3.9	-4.6	9.5	-6.4	2.2
SWE	7.5	-23.1		-13.1	-1.0	-1.1	4.7	0.7	2.7
USA	-1.8	-5.7	3.3	-11.6	0.8	-4.4	-0.5	2.8	1.5
Mean	9.5	-18.6		-14.2	-1.2	-2.1	3.6	-1.2	1.7

Counterfactual experiment of changing bilateral MP costs such that bilateral MP shares match those in 2006–2011. All numbers are in percentage points or 100× change in log points

Figure OA.1: Changes in Labor Shares and Total Inward MP Shares



Note: The figure plots the predicted change in labor shares against the change in inward MP shares. The blue dots represent the less capital-abundant countries (aggregate K/L below median) and red triangles represent the more capital-abundant ones (aggregate K/L above median).

OA.4.7 Uniform Reduction in Inward MP Costs

In this subsection, I present alternative estimates of changes in MP costs using simpler but stronger identification assumptions. Instead of estimating changes in bilateral MP costs $\hat{\gamma}_{il}$, I assume the change in MP costs of a particular host country l are the same for all potential home countries,

$$\frac{\gamma_{il}^{'}}{\gamma_{il}} = \hat{\gamma}_l^D \text{ for all } i \neq l. \tag{28}$$

I adjust $\hat{\gamma}_l^D$ to match the new total inward MP shares for 23 countries with MP data in the later period. For the 14 countries without new MP data, I simply assume the decline in their MP costs $\log \hat{\gamma}_l$ is simply equal to the global average.

Column 3 in Table OA.25 presents the calibrated change in MP costs for the 23 countries that I consider. The average decline in MP costs is 8.1 log points. As expected, countries with the largest increase in MP shares are estimated to have experienced the largest declines in their MP costs. Developed countries such as Germany, Japan and the United States have little changes in their inward MP shares and, correspondingly, the calibration shows their MP costs barely changed over this period.

In columns 4, 6, 7 and 8 of Table OA.25, I report the predicted change in labor shares, real rental rates, real wages and real income per capita. The results here are similar to those in Section 5.1 of the paper, but they also differ in some important dimensions.

Regarding the focus of the paper, the two counterfactual exercises predict very similar changes in the labor shares. The predicted changes for each country are highly correlated, with a correlation coefficient of 0.974. There is slightly more discrepancy in real wages and real rental rates between the two counterfactuals, with correlation coefficients of 0.895 and 0.879.

The biggest discrepancy occurs in the predicted change in real income per capita. The predicted changes in real income per capita only have a correlation of 0.182. In some extreme cases, such as the Netherland, the predicted changes in the two models can be of opposite signs and both have large magnitudes. To understand the causes of such discrepancies, I first establish the relationship between the change in net MP shares and welfare changes, where the net MP share of country i is the total outward MP sales net of total inward MP sales scaled by total output

$$\frac{\sum_{l \neq i} X_{il} - \sum_{j \neq i} X_{ji}}{Y_i}. \tag{29}$$

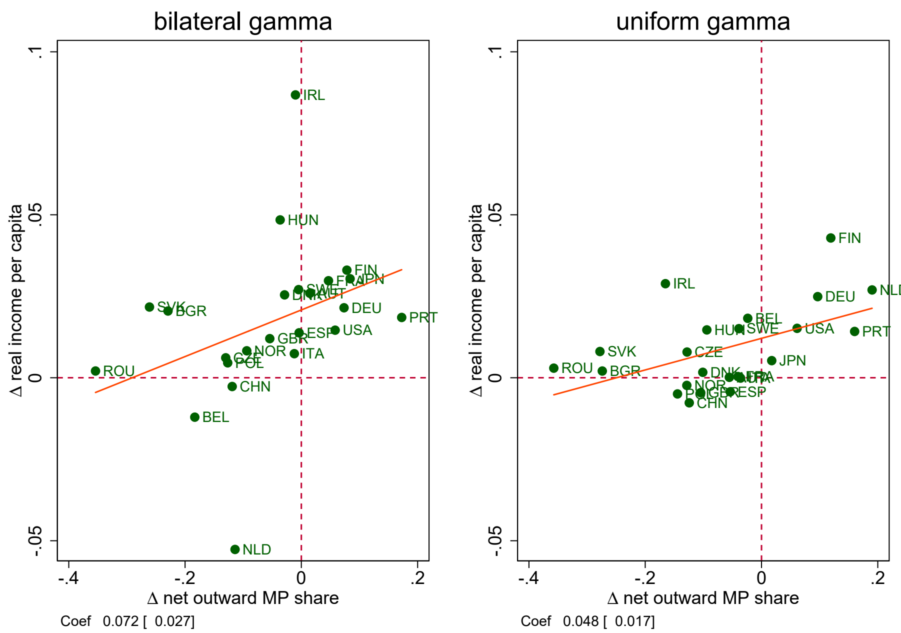
It is clear from Figure OA.2 that net outward MP share is positively associated with the change in welfare, both in the counterfactual in Section 5.1 (left panel, "bilateral gamma") and in the counterfactual considered here (right panel, "uniform gamma"). Arkolakis et al. (2018) show analytically that an important component of the gains from openness is the net

Table OA.25: Counterfactual with a Uniform Reduction in Inward MP Costs

ISO3	(1) inward MP share 96–01	(2) inward MP share 06–11	(3) Δ log MP costs (calibrated)	(4) Δ labor share (model)	(5) Δ labor share (data)	(6) $\Delta \log(r/P)$	(7) $\Delta \log(w/P)$	(8) Δ log real income
AUT	25.6	31.8	-2.0	-0.1	-2.5	0.2	-0.1	0.0
BEL	40.8	46.5	-2.1	-0.1	-2.1	2.0	1.7	1.8
BGR	3.5	31.8	-34.4	-7.0	-7.5	12.1	-16.9	0.2
CHN	2.4	15.4	-27.1	-2.5	-4.6	3.5	-7.1	-0.8
CZE	30.3	45.7	-6.5	-1.9	-0.8	4.5	-3.2	0.8
DEU	23.7	22.4	-0.2	0.8	-5.1	0.7	3.8	2.5
DNK	11.7	23.8	-10.9	-1.0	2.1	2.5	-1.6	0.2
ESP	15.2	20.9	-4.0	-0.2	-0.3	0.1	-0.8	-0.4
FIN	17.1	20.0	-5.4	0.8	3.3	2.3	5.6	4.3
FRA	16.1	21.5	-4.0	0.2	1.2	-0.4	0.3	0.0
GBR	26.6	35.3	-3.9	-0.5	0.0	0.6	-1.4	-0.5
HUN	39.6	50.5	-4.0	-1.8	-4.6	4.6	-2.7	1.5
IRL	36.0	51.4	-8.6	-3.5	5.8	8.8	-5.5	2.9
ITA	11.0	16.5	-6.3	0.3	3.1	-0.7	0.4	-0.0
JPN	3.9	4.0	0.7	0.4	-2.1	-0.3	1.1	0.5
NLD	34.6	31.6	1.4	0.9	-1.3	0.8	4.3	2.7
NOR	11.0	25.1	-14.0	-1.4	-6.4	2.8	-2.7	-0.2
POL	19.9	35.3	-8.7	-2.0	-13.3	3.3	-4.7	-0.5
PRT	33.9	20.1	9.0	1.9	1.2	-2.5	5.0	1.4
ROU	5.6	42.4	-34.6	-8.2	-4.1	14.0	-20.2	0.3
SVK	20.0	49.9	-16.3	-5.3	-4.6	10.7	-10.9	0.8
SWE	24.8	32.4	-4.6	-0.8	-1.1	3.0	-0.1	1.5
USA	12.6	10.8	1.2	0.5	-4.4	0.3	2.4	1.5
Mean	20.3	29.8	-8.1	-1.3	-2.1	3.2	-2.3	0.9

Counterfactual experiment of changing inward MP costs such that inward MP shares match those in 2006–2011. All numbers are in percentage points or 100× change in log points

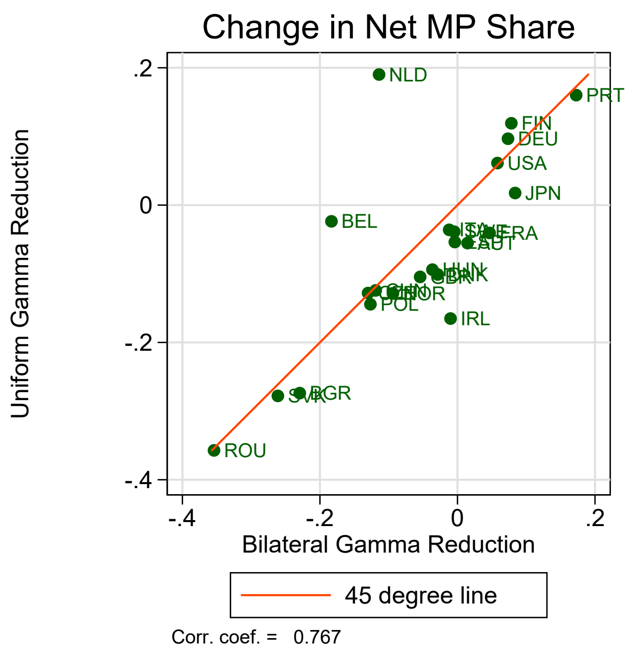
Figure OA.2: Changes in Real Income Per Capita and Net Outward MP Shares



Note: I plot the change in log real income per capita against the change in net outward MP share in both panels. Each dot represents a country. The left panel (bilateral gamma) plots the results predicted by the counterfactual in Section 5.1 of the paper, where I estimate changes in MP costs by matching bilateral MP shares. The right panel uses results predicted by the counterfactual described here, where I assume there is a uniform reduction in MP costs against all potential home countries for each host country.

flow of profits due to MP in their model, which is also the case here.

Figure OA.3: Changes in Net Outward MP Shares in the two Counterfactuals



Note: This figure plots the predicted change in net outward MP share in the counterfactual described here against that obtained in the counterfactual described in Section 5.1 of the paper.

Figure OA.3 plots the changes in net MP shares defined in equation (29) for each country in the two counterfactuals. The countries that deviate the most from the 45 degree line are the Netherlands, Belgium and Ireland. This is the main reason why the predictions on real income are so different between the two counterfactuals for these three countries. For example, the counterfactual in Section 5.1 predicts that the Netherlands has a 11-percentage-point reduction in net outward MP share and a 5.2-log-point reduction in real income, while the counterfactual here predicts that it has a 19-percentage-point increase in net outward MP share and a 2.7-log-point increase in real income. A closer look at the bilateral MP shares originated from the Netherlands reveals that most of the discrepancy in the net outward MP share is driven by two host countries, the US and Germany. In the data, firms from the Netherlands contribute to 0.9% and 2.3% of the total output in these two countries, respectively. However, the counterfactual here does not match the bilateral MP shares and predicts these two shares to be 1.3% and 5.7%. Though the differences may seem small, they strongly inflate the outward MP shares of the Netherlands because it is a much smaller economy than the US and Germany. This shows the importance of matching the bilateral MP shares, which can have large effects on aggregate variables in relatively small economies. I therefore prefer the counterfactual in Section 5.1 instead of the one considered here.

OA.4.8 Other Factors that Affect the Labor Share

Table OA.26 shows the change in endowment in each country during the period of study, as well as the calibrated change in trade costs, MP costs and home-country specific capital shifters.

After obtain all the calibrated changes, I feed them into the model one at a time to decompose their impact. The effect of each component depends on the order when adding them to the model. I consider the following four orders:

1. trade costs – MP costs – endowment – technology (see Table OA.27),
2. MP costs – trade costs – endowment – technology (see Table OA.28),
3. endowment – trade costs – MP costs – technology (see Table OA.29),
4. endowment – MP costs – trade costs – technology (see Table OA.30).

Table OA.26: Summary of (estimated) Changes in Model Primitives

ISO3	(1) $\Delta \log(K_i/L_i)$	(2) $\overline{\log(\hat{\tau}_{il})}\Big\vert_l$	(3) $\overline{\log(\hat{\tau}_{il})}\Big\vert_i$	(4) $\overline{\log(\hat{\gamma}_{il})}\Big\vert_l$	(5) $\overline{\log(\hat{\gamma}_{il})}\Big\vert_i$	(6) $\lambda_i' - \lambda$
ARG	8.4	-32.6	-2.2	-10.6^i	-25.1^i	0.006
AUS	22.8	-28.2	0.4	-11.9^i	-7.0	-0.012
AUT	11.9	-8.3	-23.6	-25.2	-27.1	0.058
BEL	19.0	-45.6	19.4	-3.2^i	-2.8	-0.001
BGR	41.8	-6.1	-41.1	-36.2	-46.0^i	0.086
BRA	-2.7	11.4	-24.1	-11.2^i	1.7	-0.096
CAN	25.7	-51.3	35.9	-11.1^i	5.7	-0.047
CHE	12.5	-17.9	-8.4	-13.5^i	-1.3	-0.038
CHN	94.5	19.1	-71.4	4.8^i	-43.6	0.170
CZE	31.8	7.6	-59.6	-39.0	-22.5	0.004
DEU	5.7	-17.8	-5.6	-6.9	-18.7	0.094
DNK	21.9	-15.4	-4.7	-31.1	-5.4	-0.022
DOM	34.9	-88.8	65.4	-11.8^i	29.6^i	-0.051
ESP	19.2	-34.5	9.8	-20.1	-17.3	0.033
FIN	13.8	-20.0	1.6	-21.6	0.7	-0.037
FRA	13.5	-28.8	8.1	-25.4	-11.1	-0.001
GBR	24.2	-48.5	29.3	-32.4	-21.6	0.032
GRC	21.2	-22.8	14.0	-12.6^i	-14.5	0.014
HUN	33.9	6.7	-54.5	-15.5	-47.6^i	0.114
IND	59.1	-81.0	26.6	-11.9^i	-40.6	0.136
IRL	67.9	-53.1	38.0	-32.0	15.2	-0.083
ITA	13.9	-41.7	30.1	-21.2	4.2	-0.040
JPN	21.5	-12.9	-11.8	-9.5^i	-29.1	0.077
KOR	47.7	-45.8	16.1	-12.6^i	-16.5	0.010
MEX	16.4	-56.6	6.2	-9.8^i	-30.4^i	0.090
NLD	16.7	-22.0	5.3	-8.7	-11.7	0.060
NOR	19.3	22.3	-56.6	-19.4	-39.6	0.130
POL	34.1	30.2	-92.0	-1.5	-65.2	0.305
PRT	40.8	-5.9	-2.3	-0.3	-20.7	0.068
ROU	69.7	11.9	-62.5	-33.6	-17.8^i	0.044
RUS	3.7	86.0	-107.0	-13.3^i	-12.0	-0.091
SVK	34.6	8.5	-81.7	-29.8	-32.3^i	0.062
SWE	19.5	-35.8	16.7	-26.2	-13.7	0.022
TUR	46.4	-6.3	-33.1	-11.7^i	-34.0	0.042
URY	1.3	-49.7	10.5	-11.5^i	-28.2^i	0.137
USA	23.9	-10.6	-4.0	4.2	-24.2	0.116
VEN	3.0	116.8	-154.6	-11.7^i	-35.9^i	-0.013
Mean	26.9	-15.3	-15.3	-16.1	-19.1	0.037

Column (1) reports the changes in endowments observed in the data. Feeding this into the model, I jointly estimate the changes in bilateral trade costs $\hat{\tau}_{il}$, MP costs $\hat{\gamma}_{il}$ and changes in home-country capital share shifter $\lambda_i' - \lambda$ by matching the changes in bilateral trade shares, total inward MP shares and the labor shares from the base period to 2006–2011. All numbers are in percentage points or $100\times$ change in log points except for those in column (6). Statistics in columns (2) - (5) are the average change in trade costs by destination, average change in trade costs by origin, average change in MP costs by destination, and average change in MP costs by origin, respectively. Superscript i in columns (4) and (5) indicates averages calculated based on fewer than 10 observations.

Table OA.27: Step-by-step Decomposition of Changes in the Labor Share

Country Code	τ	γ	(K_i, L_i)	λ_i	Total
AUT	0.1	-0.5	0.6	-2.7	-2.5
BEL	3.0	-7.7	2.0	0.5	-2.1
BGR	-0.2	-3.7	2.4	-5.9	-7.5
CHN	-0.1	-1.2	7.6	-10.9	-4.6
CZE	2.0	-6.2	2.6	0.7	-0.8
DEU	-0.6	1.7	0.2	-6.3	-5.0
DNK	-0.2	-1.9	1.1	3.0	2.1
ESP	-0.0	-0.4	2.0	-1.9	-0.3
FIN	-0.4	-2.7	0.9	5.5	3.3
FRA	-0.2	-0.4	0.8	0.9	1.2
GBR	0.3	-1.2	2.0	-1.1	0.0
HUN	4.0	-3.4	1.9	-7.0	-4.6
IRL	3.4	-15.1	10.1	7.2	5.7
ITA	0.0	-0.5	0.9	2.8	3.1
JPN	-0.2	1.0	1.4	-4.3	-2.1
NLD	0.2	0.3	0.9	-2.7	-1.3
NOR	-0.2	1.1	1.4	-8.7	-6.4
POL	0.6	0.8	2.4	-17.0	-13.2
PRT	-1.2	3.9	2.8	-4.3	1.2
ROU	-0.1	-4.7	2.6	-1.8	-4.1
SVK	2.3	-5.3	2.1	-3.8	-4.6
SWE	0.5	-3.2	1.2	0.4	-1.1
USA	-0.4	1.6	1.9	-7.5	-4.4
Avg	0.6	-2.1	2.2	-2.8	-2.1

The above table shows the percentage point change in labor shares by introducing changes in trade costs τ, MP costs γ, endowments (K_i, L_i) and country-specific technologies λ_i one at a time. Changes in endowments are obtained from the data, while changes in the other three model primitives are jointly calibrated by targeting the changes in trade shares, MP shares and labor shares. Each column shows the additional impact on the labor share after introducing the change. The four components add up to the total effect on labor shares (last column), which, by construction, is the same as the change in labor shares in the data.

Table OA.28: Step-by-step Decomposition of Changes in the Labor Share

Country Code	γ	τ	(K_i, L_i)	λ_i	Total
AUT	-0.6	0.2	0.6	-2.7	-2.5
BEL	-6.6	1.9	2.0	0.5	-2.1
BGR	-6.9	3.0	2.4	-5.9	-7.5
CHN	-1.8	0.5	7.6	-10.9	-4.6
CZE	-6.8	2.7	2.6	0.7	-0.8
DEU	1.4	-0.3	0.2	-6.3	-5.0
DNK	-2.2	0.1	1.1	3.0	2.1
ESP	-0.4	-0.0	2.0	-1.9	-0.3
FIN	-3.2	0.1	0.9	5.5	3.3
FRA	-0.3	-0.2	0.8	0.9	1.2
GBR	-1.5	0.6	2.0	-1.1	0.0
HUN	-3.7	4.3	1.9	-7.0	-4.6
IRL	-11.0	-0.7	10.1	7.2	5.7
ITA	-0.4	-0.1	0.9	2.8	3.1
JPN	1.0	-0.2	1.4	-4.3	-2.1
NLD	0.4	0.2	0.9	-2.7	-1.3
NOR	3.5	-2.6	1.4	-8.7	-6.4
POL	1.4	0.0	2.4	-17.0	-13.2
PRT	3.1	-0.4	2.8	-4.3	1.2
ROU	-7.6	2.8	2.6	-1.8	-4.1
SVK	-9.1	6.1	2.1	-3.8	-4.6
SWE	-3.7	1.0	1.2	0.4	-1.1
USA	1.4	-0.1	1.9	-7.5	-4.4
Avg	-2.3	0.8	2.2	-2.8	-2.1

The above table shows the percentage point change in labor shares by introducing changes in trade costs τ, MP costs γ, endowments (K_i, L_i) and country-specific technologies λ_i one at a time. Changes in endowments are obtained from the data, while changes in the other three model primitives are jointly calibrated by targeting the changes in trade shares, MP shares and labor shares. Each column shows the additional impact on the labor share after introducing the change. The four components add up to the total effect on labor shares (last column), which, by construction, is the same as the change in labor shares in the data.

Table OA.29: Step-by-step Decomposition of Changes in the Labor Share

Country Code	(K_i, L_i)	τ	γ	λ_i	Total
AUT	0.8	0.1	-0.6	-2.7	-2.5
BEL	1.7	3.1	-7.4	0.5	-2.1
BGR	2.7	-0.1	-4.2	-5.9	-7.5
CHN	7.1	-0.2	-0.6	-10.9	-4.6
CZE	2.2	2.0	-5.8	0.7	-0.8
DEU	0.0	-0.6	1.9	-6.3	-5.0
DNK	1.2	-0.1	-2.1	3.0	2.1
ESP	1.8	-0.1	-0.1	-1.9	-0.3
FIN	0.9	-0.4	-2.6	5.5	3.3
FRA	0.8	-0.2	-0.4	0.9	1.2
GBR	2.0	0.2	-1.1	-1.1	0.0
HUN	1.7	4.4	-3.6	-7.0	-4.6
IRL	6.3	3.5	-11.4	7.2	5.7
ITA	0.9	0.0	-0.6	2.8	3.1
JPN	1.3	-0.2	1.0	-4.3	-2.1
NLD	1.1	0.2	0.1	-2.7	-1.3
NOR	1.6	-0.3	1.0	-8.7	-6.4
POL	2.3	0.6	0.8	-17.0	-13.2
PRT	3.0	-0.9	3.4	-4.3	1.2
ROU	3.6	0.6	-6.5	-1.8	-4.1
SVK	1.8	2.9	-5.5	-3.8	-4.6
SWE	1.4	0.5	-3.4	0.4	-1.1
USA	1.7	-0.4	1.8	-7.5	-4.4
Avg	2.1	0.7	-2.0	-2.8	-2.1

The above table shows the percentage point change in labor shares by introducing changes in trade costs τ, MP costs γ, endowments (K_i, L_i) and country-specific technologies λ_i one at a time. Changes in endowments are obtained from the data, while changes in the other three model primitives are jointly calibrated by targeting the changes in trade shares, MP shares and labor shares. Each column shows the additional impact on the labor share after introducing the change. The four components add up to the total effect on labor shares (last column), which, by construction, is the same as the change in labor shares in the data.

Table OA.30: Step-by-step Decomposition of Changes in the Labor Share

Country Code	(K_i, L_i)	γ	τ	λ_i	Total
AUT	0.8	-0.9	0.3	-2.7	-2.5
BEL	1.7	-6.5	2.1	0.5	-2.1
BGR	2.7	-7.3	3.0	-5.9	-7.5
CHN	7.1	-1.0	0.2	-10.9	-4.6
CZE	2.2	-5.9	2.2	0.7	-0.8
DEU	0.0	1.4	-0.2	-6.3	-5.0
DNK	1.2	-2.4	0.2	3.0	2.1
ESP	1.8	-0.2	-0.1	-1.9	-0.3
FIN	0.9	-3.3	0.3	5.5	3.3
FRA	0.8	-0.5	-0.1	0.9	1.2
GBR	2.0	-1.4	0.6	-1.1	0.0
HUN	1.7	-3.6	4.4	-7.0	-4.6
IRL	6.3	-7.0	-0.8	7.2	5.7
ITA	0.9	-0.5	-0.1	2.8	3.1
JPN	1.3	1.1	-0.2	-4.3	-2.1
NLD	1.1	0.0	0.3	-2.7	-1.3
NOR	1.6	3.1	-2.3	-8.7	-6.4
POL	2.3	1.7	-0.2	-17.0	-13.2
PRT	3.0	2.9	-0.4	-4.3	1.2
ROU	3.6	-8.7	2.9	-1.8	-4.1
SVK	1.8	-8.2	5.6	-3.8	-4.6
SWE	1.4	-3.9	1.0	0.4	-1.1
USA	1.7	1.6	-0.2	-7.5	-4.4
Avg	2.1	-2.2	0.8	-2.8	-2.1

The above table shows the percentage point change in labor shares by introducing changes in trade costs τ, MP costs γ, endowments (K_i, L_i) and country-specific technologies λ_i one at a time. Changes in endowments are obtained from the data, while changes in the other three model primitives are jointly calibrated by targeting the changes in trade shares, MP shares and labor shares. Each column shows the additional impact on the labor share after introducing the change. The four components add up to the total effect on labor shares (last column), which, by construction, is the same as the change in labor shares in the data.

OA.5 Calibration and Counterfactuals under Alternative Setups

In this section, I present calibration and counterfactual results for three alternative setups of the model.

OA.5.1 Model without endogenous technology choice

As discussed in the paper, I can shut down endogenous technology choice by setting $\eta \to -\infty$. In this scenario, all firms choose the same technology $(a, b) = (1, 1)$. I then recalibrate the model by targeting the same set of parameters except the technology origin effect. Table OA.31 shows the calibrated outer loop parameters. I then perform the first counterfactual experiment by calibrating the change in MP costs in the decade after 1996–2001. The results are presented in Table OA.32.

Table OA.31: Calibration - Targets and Parameters - TCC only

Parameters	Values/Normalization	Targets
ξ	0.593	coefficient of revenue 0.054
k	4.275	unrestricted trade elasticity 4.3
χ	11.001	restricted trade elasticity 10.9
λ	0.350	average labor share 0.520

Table OA.32: Counterfactual of Reduction in Bilateral MP Costs: Technology-Capital Complementarity Only

ISO3	(1) Δinward MP share	(2) $\overline{\log(\hat{\gamma}_{il})}$ for each l	(3) Δoutward MP share	(4) $\overline{\log(\hat{\gamma}_{il})}$ for each i	(5) Δlabor share (model)	(6) Δlabor share (data)	(7) $\Delta \log(r/P)$	(8) $\Delta \log(w/P)$	Δlog real income
AUT	6.2	-25.7		-20.1	-0.9	-2.5	5.1	1.4	2.7
BEL	5.7	16.8^{i}		-6.6	-1.5	-2.1	2.8	-3.7	-1.1
BGR	28.3	-35.2		-29.0^{i}	-2.8	-7.5	5.0	-8.7	1.1
CHN	13.0	-0.8^{i}		-21.7	-0.0	-4.6	-0.4	-0.7	-0.5
CZE	15.3	-22.6		-17.2	-2.1	-0.8	4.4	-4.2	0.3
DEU	-1.2	-11.4	3.7	-9.9	0.0	-5.1	2.1	2.2	2.2
DNK	12.2	-28.8		-7.2	-1.2	2.1	5.8	0.8	2.7
ESP	5.7	-15.2	6.2	-2.6	-0.2	-0.3	1.9	1.1	1.4
FIN	2.9	-16.4	10.5	-11.6	-0.7	3.3	5.4	2.3	3.3
FRA	5.4	-21.9	15.1	-14.5	-0.0	1.2	3.0	2.9	2.9
GBR	8.7	-35.8	6.7	-19.6	-0.3	0.0	1.9	0.6	1.2
HUN	10.9	-15.8	7.3	-27.2^{i}	-1.1	-4.6	5.8	1.4	4.0
IRL	15.3	-23.7		-12.7	-2.9	5.8	12.9	0.8	7.9
ITA	5.5	-12.9	5.1	-13.9	-0.2	3.1	1.4	0.5	0.8
JPN	0.1	-15.4^{i}	7.3	-22.4	0.0	-2.1	3.0	3.1	3.1
NLD	-3.0	-15.4		-7.6	0.6	-1.3	-6.8	-4.1	-5.2
NOR	14.1	-26.9	4.6	-11.7	-1.9	-6.4	5.9	-2.1	1.0
POL	15.4	-27.6		-22.5	-1.2	-13.3	2.3	-2.6	0.1
PRT	-13.7	-10.3	4.6	-6.1	1.6	1.2	-1.6	4.8	1.9
ROU	36.9	-29.9		-0.3^{i}	-2.6	-4.1	2.2	-10.6	-1.5
SVK	29.9	-14.5		-18.8^{i}	-3.0	-4.6	6.8	-5.4	1.5
SWE	7.5	-23.1		-12.7	-0.9	-1.1	4.4	0.8	2.6
USA	-1.8	-5.8	3.3	-11.6	0.1	-4.4	1.4	1.9	1.7
Mean	9.5	-18.2		-14.2	-0.9	-2.1	3.3	-0.8	1.5

Counterfactual experiment of changing bilateral MP costs such that MP shares match those in 2006–2011. All numbers are in percentage points or 100× change in log points

OA.5.2 Exogenous technology differences

In this subsection, I consider an alternative setup where firms' technologies are exogenously given and differ based on their home countries. In particular, I assume the capital share parameter λ is home-country specific and it increases with the home country's capital abundance in a log linear way

$$\log\left(\frac{\lambda_i}{1-\lambda_i}\right) = \alpha_0^\lambda + \alpha_1^\lambda \log\left(K_i/L_i\right). \tag{30}$$

A foreign affiliate from country i inherits λ_i from its parent firm. Therefore, the extensive elasticity η is replaced by a new parameter α_1^λ, which controls the extent to which the capital intensities of the technologies correlates with endowments in the home countries. I recalibrate the model targeting exactly the same moments as in the baseline model. Table OA.33 presents the calibrated parameters, and Table OA.34 shows the results for the first counterfactual experiment.

Table OA.33: Calibration - Targets and Parameters - Exogenous TOE

Parameters	Values/Normalization	Targets
ξ	0.603	coefficient of revenue 0.054
k	4.204	unrestricted trade elasticity 4.3
χ	10.940	restricted trade elasticity 10.9
α_0^λ	-0.513	average labor share 0.520
α_1^λ	0.269	technology origin effect 0.277

Table OA.34: Counterfactual of Reduction in Bilateral MP Costs: Exogenous Technology Differences

ISO3	(1) inward MP share 96–01	(2) inward MP share 06–11	(3) $\overline{\log(\hat{\gamma}_{il})}$ for each l	(4) $\overline{\log(\hat{\gamma}_{il})}$ for each i	(5) Δ labor share (model)	(6) Δ labor share (data)	(7) $\Delta log(r/P)$	(8) $\Delta log(w/P)$	Δ log real income
AUT	6.2	-25.2		-20.8	-0.4	-2.5	3.7	1.9	2.7
BEL	5.7	17.7^i		-7.0	-1.9	-2.1	2.8	-4.7	-1.3
BGR	28.3	-39.2		-28.7^i	-7.1	-7.5	13.7	-16.4	2.1
CHN	13.0	-5.1^i		-21.7	-2.9	-4.6	4.4	-8.0	-0.2
CZE	15.3	-23.0		-18.0	-2.5	-0.8	5.4	-4.9	0.6
DEU	-1.2	-11.3	3.7	-10.7	0.2	-5.1	1.8	2.5	2.2
DNK	12.2	-28.6		-7.6	-0.8	2.1	4.3	1.2	2.5
ESP	5.7	-15.1	6.2	-3.1	-0.0	-0.3	1.4	1.3	1.4
FIN	2.9	-16.5	10.5	-12.7	-0.7	3.3	5.2	2.5	3.6
FRA	5.4	-21.8	15.1	-16.5	0.2	1.2	2.5	3.4	3.1
GBR	8.7	-36.1	6.7	-20.3	-0.1	0.0	1.4	1.0	1.2
HUN	10.9	-16.3	7.3	-27.6^i	-2.0	-4.6	7.9	-0.4	4.5
IRL	15.3	-22.8		-12.8	-3.8	5.8	14.9	-0.8	8.5
ITA	5.5	-12.3	5.1	-15.6	0.0	3.1	0.8	0.8	0.8
JPN	0.1	-14.9^i	7.3	-23.7	0.0	-2.1	3.2	3.3	3.2
NLD	-3.0	-15.3		-7.9	1.1	-1.3	-7.8	-3.4	-5.4
NOR	14.1	-26.7	4.6	-10.8	-1.7	-6.4	4.8	-2.1	0.9
POL	15.4	-28.6		-22.6	-2.3	-13.3	4.8	-4.6	0.5
PRT	-13.7	-10.3	4.6	-5.8	2.2	1.2	-2.8	5.9	1.7
ROU	36.9	-34.7		1.5^i	-8.6	-4.1	14.1	-22.9	0.2
SVK	29.9	-16.0		-19.0^i	-4.6	-4.6	10.6	-8.2	2.2
SWE	7.5	-23.5		-13.0	-1.1	-1.1	4.8	0.5	2.7
USA	-1.8	-5.6	3.3	-12.3	0.2	-4.4	1.2	1.9	1.6
Mean	9.5	-18.7		-14.6	-1.6	-2.1	4.5	-2.2	1.7

Counterfactual experiment of changing inward MP costs such that inward MP shares match those in 2006–2011. All numbers are in percentage points or 100× change in log points

OA.5.3 Capital Mobility

Table OA.35 shows the calibrated outer loop parameters when I assume capital is fully mobile across countries. Table OA.36 shows the results from the same exercise in Section 5.1 using this alternative model. The average decline in labor shares across countries is reduced to 0.44 percentage points. However, there is still a negative correlation between the change in labor shares and the change in inward MP shares (correlation coefficient is -0.72).

Table OA.35: Baseline Calibration - Targets and Parameters - Capital is mobile

Parameters	Values/Normalization	Targets
η	0.813	technology origin effect 0.277
ξ	0.598	coefficient of revenue 0.054
k	4.189	unrestricted trade elasticity 4.3
χ	10.934	restricted trade elasticity 10.9
λ	0.263	average labor share 0.520

Table OA.36: Counterfactual of Reduction in Bilateral MP Costs: Capital is Mobile

ISO3	(1) inward MP share 96–01	(2) inward MP share 06–11	(3) $\overline{\log(\hat{\gamma}_{il})}$ for each l	(4) $\overline{\log(\hat{\gamma}_{il})}$ for each i	(5) Δ labor share (model)	(6) Δ labor share (data)	(7) $\Delta log(r/P)$	(8) $\Delta log(w/P)$	Δ log real income
AUT	6.2	-23.5		-20.0	0.2	-2.5	2.2	3.0	2.6
BEL	5.7	6.4^{i}		-6.5	-1.1	-2.1	1.2	-3.4	-1.3
BGR	28.3	-35.0		-30.2^{i}	-2.3	-7.5	6.1	-3.5	2.2
CHN	13.0	-19.1^{i}		-22.7	-2.4	-4.6	4.3	-5.3	-0.0
CZE	15.3	-21.0		-18.1	-1.1	-0.8	2.4	-1.9	0.6
DEU	-1.2	-9.6	3.7	-6.7	0.2	-5.1	0.5	1.5	1.1
DNK	12.2	-27.5		-8.3	0.0	2.1	2.8	2.9	2.9
ESP	5.7	-13.4	6.2	-24.9	0.1	-0.3	1.3	1.8	1.6
FIN	2.9	-15.3	10.5	-8.2	0.9	3.3	2.0	5.6	4.2
FRA	5.4	-21.0	15.1	-19.2	1.0	1.2	1.7	5.7	3.9
GBR	8.7	-35.8	6.7	-19.6	-0.0	0.0	2.1	2.0	2.1
HUN	10.9	-14.7	7.3	-26.7^{i}	0.4	-4.6	4.3	5.8	4.9
IRL	15.3	-23.8		-10.2	1.3	5.8	9.0	15.2	13.3
ITA	5.5	-11.6	5.1	-17.5	0.2	3.1	0.5	1.3	0.9
JPN	0.1	-21.5^{i}	7.3	-14.0	1.0	-2.1	-0.7	3.5	1.8
NLD	-3.0	-15.3		-8.8	-1.2	-1.3	-2.1	-7.1	-5.0
NOR	14.1	-25.5	4.6	-12.0	-0.7	-6.4	2.9	-0.0	1.2
POL	15.4	-26.3		-23.0	-1.2	-13.3	2.7	-2.3	0.5
PRT	-13.7	-8.8	4.6	-5.6	0.5	1.2	0.8	2.6	1.6
ROU	36.9	-33.1		-3.5^{i}	-4.6	-4.1	9.6	-9.6	1.9
SVK	29.9	-14.1		-18.8^{i}	-2.2	-4.6	6.4	-2.6	2.5
SWE	7.5	-24.1		-12.0	0.3	-1.1	3.5	4.7	4.2
USA	-1.8	-4.6	3.3	-7.5	0.5	-4.4	-0.1	1.8	1.0
Mean	9.5	-19.0		-15.0	-0.4	-2.1	2.7	0.9	2.1

Counterfactual experiment of changing inward MP costs such that inward MP shares match those in 2006–2011. All numbers are in percentage points or 100× change in log points

OA.5.4 Marketing/Entry Costs Paid in Factors

The recent quantitative MP models such as ARRY assume that marketing and entry costs are paid using factors such as labor. In the baseline model, I deviate from the literature and assume that marketing and entry costs are paid using composite goods. The reason for this deviation is that, with two factors, it is unclear how capital and labor are mixed for the marketing and entry activities. Without more detailed data, it is hard to estimate the capital intensity of these activities. Here I provide some robustness checks by assuming the marketing and entry costs are paid using both capital and labor while setting somewhat arbitrary capital intensities for these two activities. Specifically, I will assume that firms pay the marketing costs by combining capital and labor in the destination country in a Cobb-Douglas fashion, and the labor share is the same as the country-level labor share in the baseline period (1996-2001). Similarly, the firms pay the entry costs by combining capital and labor in the home country, with the labor share being the same as the aggregate.[8]

I recalibrate the model by targeting the same moments as in the paper and the results are presented in Table OA.37. There is little change to the calibrated values of k and χ, while the parameters controlling the extensive substitution and technology-capital complementarity become larger. A significant difference between this calibration and the baseline is that the aggregate prices here are much higher. This is because, in the baseline model, marketing and entry costs are paid by the composite goods. Finally, the predicted labor share decline, which is calculated in the same way as that in Section 5.1 of the paper, is slightly smaller in this alternative calibration.

Table OA.37: Baseline Calibration - Targets and Parameters - Marketing and Entry Costs Paid in Factors

Parameters	Values/Normalization	Targets
η	0.850	technology origin effect 0.277
ξ	0.835	coefficient of revenue 0.054
k	4.069	unrestricted trade elasticity 4.3
χ	10.937	restricted trade elasticity 10.9
λ	0.227	average labor share 0.520

[8] We do not vary the capital intensities when we perform the counterfactuals because it is unclear how they would evolve over time. Given the importance of intangible capital, it is valuable to better quantify the capital intensities of marketing and entry activities using the right data moments. I leave it for future work.

Table OA.38: Compare Calibration to Baseline

	$\overline{\log P}$	$\overline{\log M}$	$\overline{\log F_e}$	$\overline{\log(r/w)}$	$\Delta\%$ Labor Share
Baseline	3.15	-8.76	0.99	0.44	-1.18
Marketing/Entry using Factors	4.98	-5.87	1.18	0.44	-0.96

Average change in labor shares (for 23 countries with MP data in 2006-2011) is the result from the counterfactual exercises described in Section 5.1 of the paper. Changes in MP costs are recalibrated under the new parameters.

OA.5.5 Different Strategies to Calibrate Entry Costs

As I discuss in the paper, it is unclear how one should discipline the entry costs F_{ei} in the model. In the baseline calibration, I find F_{ei} such that the fraction of entrants that end up serving the domestic markets is 70% in all countries. In this subsection, I consider three alternative assumptions:

1. the probability of serving domestic market is 70% for US firms, and the entry costs are the same across countries (so the probability of serving domestic markets may not be 70% for firms from other countries)
2. the probability of serving domestic markets is 10% in all countries
3. the probability of serving domestic market is 10% for US firms and the entry costs are the same across countries.

The calibration and counterfactual results under these three alternative assumptions, together with the baseline calibration, are presented in Table OA.39. The calibrated parameters are similar across different columns. Moreover, regarding the counterfactual exercise in Section 5.1, the average labor share decline due to the change in MP costs is 1.2 percentage points in all calibrations. The correlation between the labor share decline in each country is almost perfect across different calibrations. A similar pattern occurs for real income per capita. In total, the assumptions used to pin down the entry costs F_{ei} are not crucial regarding the most important predictions of the model.

Table OA.39: Compare Calibration to Baseline: Alternative Assumption on Entry Costs

	$\Pr(\phi \geq \phi_{ii}) = 0.7, \forall i$	$\Pr(\phi \geq \phi_{US,US}) = 0.7$ $F_{ei} = F_{e,US}, \forall i$	$\Pr(\phi \geq \phi_{ii}) = 0.1, \forall i$	$\Pr(\phi \geq \phi_{US,US}) = 0.1$ $F_{ei} = F_{e,US}, \forall i$
η	0.545	0.410	0.553	0.391
Implied Total Elas	0.764	0.732	0.766	0.728
ξ	0.580	0.591	0.555	0.588
k	4.201	4.199	4.206	4.200
χ	10.932	10.941	10.928	10.942
λ	0.313	0.318	0.289	0.292
Average Δ% Labor Share	-1.2	-1.2	-1.2	-1.2
Correlation with Baseline		0.999	1.000	0.999
Average Δ% log income per capita	1.7	1.7	1.7	1.7
Correlation with Baseline		1.000	1.000	1.000

Average change in labor shares (for 23 countries with MP data in 2006-2011) is the result from the counterfactual exercises described in Section 5.1 of the paper. Changes in MP costs are recalibrated under the new parameters.

References

Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple, "Innovation and Production in the Global Economy," *American Economic Review*, August 2018, *108* (8), 2128–2173.

Atalay, Enghin, Ali Hortaçsu, and Chad Syverson, "Vertical Integration and Input Flows," *American Economic Review*, April 2014, *104* (4), 1120–1148.

Burstein, Ariel and Jonathan Vogel, "International Trade, Technology, and the Skill Premium," *Journal of Political Economy*, June 2017, *125* (5), 1356–1412.

Crozet, Matthieu and Federico Trionfetti, "Firm-Level Comparative Advantage," *Journal of International Economics*, November 2013, *91* (2), 321–328.

Ethier, Wilfred J. and James R. Markusen, "Multinational firms, technology diffusion and trade," *Journal of International Economics*, August 1996, *41* (1), 1–28.

Gravelle, Jane G., "Tax Havens: International Tax Avoidance and Evasion," *National Tax Journal*, 2009, *62* (4), 727–753.

Guillaume, Gaulier and Soledad Zignago, "BACI: International Trade Database at the Product-Level. The 1994-2007 Version," Working Papers 2010-23, CEPII Research Center 2010.

Karabarbounis, Loukas and Brent Neiman, "The Global Decline of the Labor Share," *The Quarterly Journal of Economics*, February 2014, *129* (1), 61–103.

Keller, Wolfgang and Stephen Ross Yeaple, "The Gravity of Knowledge," *American Economic Review*, 2013, *103* (4), 1414–44.

Oi, Walter Y., "The Desirability of Price Instability Under Perfect Competition," *Econometrica*, 1961, *29* (1), 58–64.

Ramondo, Natalia, Andrés Rodríguez-Clare, and Felix Tintelnot, "Multinational Production: Data and Stylized Facts," *American Economic Review, Papers and Proceedings*, 2015.

_ **, Veronica Rappoport, and Kim J. Ruhl**, "Intrafirm trade and vertical fragmentation in U.S. multinational corporations," *Journal of International Economics*, January 2016, *98*, 51–59.

Ravikumar, B., Ana Santacreu, and Michael Sposi, "Capital Accumulation and Dynamic Gains from Trade," FRB St. Louis Working Paper February 2017.

Robert Inklaar and Marcel P. Timmer, "Capital, Labor, and TFP in PWT8.0," Working Paper July 2013.