Supplementary Appendix (not for Publication)
Talent, Geography, and Offshore R&D

Jingting Fan
Pennsylvania State University
This appendix presents supplementary materials that are not essential for the main text

Contents

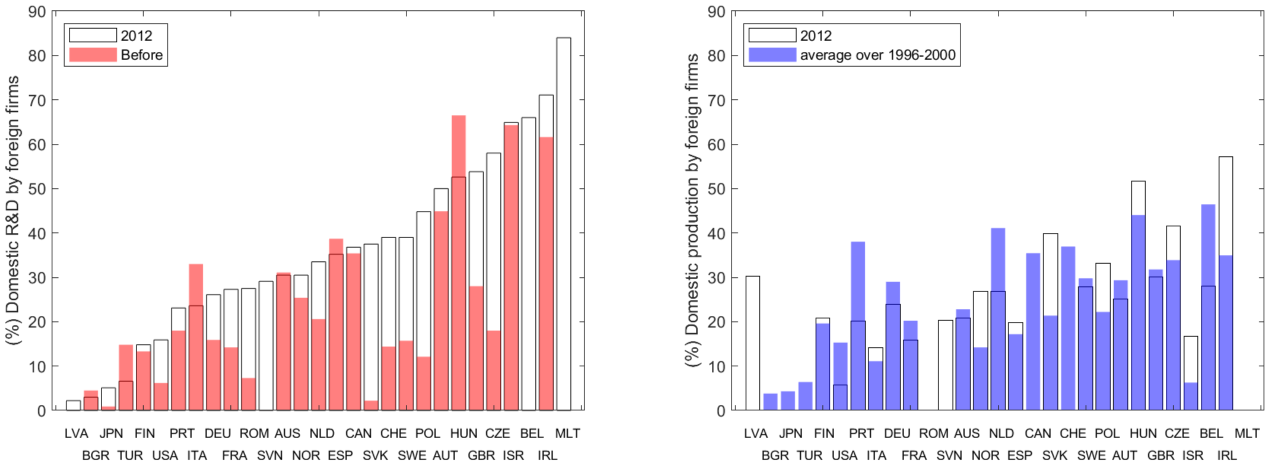
SA.A Empirics

SA.A.1 The (Growing) Importance of Offshore R&D

The left panel of Figure SA.A.1 shows that in many host countries, R&D carried out by foreign firms has increased over the decade and amounted to a big part of domestic R&D by 2012. As a comparison, the right panel of Figure SA.A.1 depicts the share of production (using sales as proxy) in a country done by foreign firms, calculated from the OECD database and the MP dataset in Ramondo et al. (2019) (see Ramondo et al., 2015 for descriptions).[1]

Three observations emerge from this comparison. First, host economies with a high share of R&D at foreign firms also tend to have a high share of production at foreign firms. The correlation between the

[1]The primary source of data for both panels in the figure is the OECD Activity of Multinational Enterprises (AMNE) database (OECD, 2015). The coverage of this database differs by years and variable. For production (or its proxies, such as value added or turnover), the coverage is incomplete until after 2010, so I use the share calculated from Ramondo et al. (2015) for the starting period (blue bars). Ramondo et al. (2015) measure MP using sales. To maintain consistency, I use turnovers from the AMNE database to measure production for 2012. Using value added or production gives essentially the same pattern.



(a) Growth in Offshore R&D

(b) Growth in Offshore Production

Figure SA.A.1: The Level and Growth of Offshore R&D (and Production)

Notes: Shown in the left panel is $\frac{\text{Business enterprise R\&D expenditures in country } i \text{ by foreign firms}}{\text{Total business enterprise R\&D expenditures in country } i}$. Uncolored bars are for 2012; colored bars are for the beginning of the sample, which differs by country and dates back to as early as 1985. The source of the data is the OECD. Shown in the right panel is the share of total sales in a country generated in affiliates of foreign firms. The source of the data is OECD and Ramondo et al. (2015).

two ratios is 0.39 in the beginning of the period and 0.55 in 2012. Second, unlike offshore R&D, offshore production did not increased dramatically over the period, especially among developed countries in this sample.[2] Lastly, after rapid growth, offshore R&D shares are somewhat higher than multinational production (MP) shares in these countries. In 2012, the median shares of foreign R&D and production being 33.5% and 25.9%, respectively.

Figure SA.A.1 shows that offshore R&D is closely connected with, but distinct from, multinational production. It is exactly this distinction and the connection that I seek to capture in the structural model.

SA.A.2 Robustness: Alternative Measures and Sample Restriction

In this subsection, I show that the three facts are robust to different measures of affiliate R&D and production, and alternative sample restrictions.

Measure of R&D. In the baseline analysis, I map patents to inventor countries based on the share of inventors from different countries. For example, if a patent has inventors in more than one country, I assign each country the fraction of inventors residing in it. As an alternative, I count each county as having the full patent—for example, if a patent is invented jointly by one person in the U.S. headquarters and one person in Canada, I count the Canadian affiliate and the U.S. headquarters as each having invented a full patent. In what follows I will call this unweighed patent counts.

Second, it is well known that patents differ vastly in their values and that the number of forward citations to a patent is a good proxy for its value—much like the number of forward citations to an academic article often indicates its importance. To adjust for patent quality, I use the number of citations received by the patents invented at an affiliate as a measure for the invention output of that affiliate.

Finally, in the baseline analysis involving the extensive margin of offshore R&D, I define a host as having an R&D center if at least one full patent is invented there, which rules out affiliates with only a small number of joint inventions with the headquarters. For robustness, I use a more liberal definition, classifying a country as hosting an R&D center as long as a positive fraction of a patent is invented there.

[2] This might appear surprising given the narrative about the rise in multinational production, but note that in most developed countries, the rise might have taken place before 2000. The developing economies in this data, such as Slovakia, Poland, and Czech Republic do see an increase in multinational production.

Measure of production. In the baseline analysis, I measure affiliate production by sales. Some of the sales are likely due to intermediate products made elsewhere. I use value added for robustness.

Restriction to manufacturing. The baseline analysis includes firms from all industries. Since the quantitative model has a focus on trade, it is more appropriate to interpret it as for manufacturing. I show all baseline estimates remain materially the same when I restrict to manufacturing firms. This also addresses the concern that firms in the service industry—in particular the industry NAICS 54 (Professional, Scientific, and Technical Services)—might generate revenue from licensing intellectual properties, which could lead to a mechanical correlation between sales revenue and R&D.

Excluding headquarters from regressions. For Fact 1, the baseline regressions include observations that are in the headquarters country of a firm. One might be concerned that the concentration of R&D in countries with talents is driven entirely by headquarters. Although this interpretation does not alter the main message that the invention intensity of a firm differs across countries with different talent endowments, I show that the results are similar if headquarters are excluded.

I now describe the results from these robustness exercises.

Fact 1 robustness. Table SA.A.1 reports the robustness results for Fact 1. I take the last two columns of Table A.8 as the baseline for intensive and extensive margin regressions, respectively. Columns 1 to 5 focus on the intensive margin measure of invention intensity. Column 1 restricts the sample to manufacturing only; Column 2 excludes headquarters from the sample; Columns 3 to 5 permute on the dependent variables, the log ratio between invention and production, by changing the measure for either invention output or production. Throughout, the human capital index is economically sizable, although the number of narrowly defined researchers are not statistically significant in some specifications.

Columns 6 to 8 are robustness for when the dependent variable is an R&D indicator. Column 6 includes only manufacturing firms; Column 7 excludes headquarters; Column 8 uses the liberal definition of R&D centers. Results from all three specifications are qualitatively similar.

Fact 2 robustness. Tables SA.A.2 and SA.A.3 report the robustness results for Fact 2. The specifications in Table SA.A.2 reproduce Columns 1, 4, 5, and 7 of Table A.9, using the same measures for R&D and production but restricting the sample to manufacturing firms. Across specifications, the point estimates are generally close to that from the baseline sample.

Table SA.A.3 reports robustness with alternative definitions of invention and production. Columns 1 and 2 show that having an liberally defined R&D center is associated with both the presence and the size of production facilities. The coefficients are close to the baseline estimates (Columns 1 and 4 of Table A.9, respectively). Columns 3 through 6 reproduce Columns 5 and 7 of Table A.9 using citations and unweighted patent counts to measure the intensive margin of invention, respectively. Finally, Columns 7 to 9 keep the baseline measure of invention but change the dependent variable to value added, which reduces the sample size substantially. The specifications correspond to Columns 4, 5, 7 of Table A.9.

Exercises here show that Fact 2 holds across different measures and sub-samples. In Appendix SA.A.3, I show that an alternative identification strategy using the variation in host R&D subsidies and number of researchers lead to qualitatively similar findings.

Fact 3 robustness. Tables SA.A.4 and SA.A.5 report additional robustness exercises for Fact 3. Tables SA.A.4 uses the baseline measures but restricts the sample to manufacturing. The coefficients are generally similar to the baseline estimates.

Table SA.A.5 uses the same sample as the baseline and changes the measure of invention and production. Columns 1 to 3 show the headquarter effect for invention is robust to the liberal definition of R&D centers and two different measures of invention. Most coefficients are broadly in line with the corresponding ones (Columns 1 and 2) in Table A.9. Columns 4 to 6, corresponding to Columns 4 and 6 of Table A.10, show the headquarter effect for production are robust when these alternative measures of R&D are used as controls. Finally, Column 7 shows that the result is similar when production is measured using value added.

Table SA.A.1: Fact 1 Robustness

Dependent var.	(1) $\ln(\frac{\text{patent}}{\text{sales}})$	(2) $\ln(\frac{\text{patent}}{\text{sales}})$	(3) $\ln(\frac{\text{unwgt. patent}}{\text{sales}})$	(4) $\ln(\frac{\text{citation}}{\text{sales}})$	(5) $\ln(\frac{\text{patent}}{\text{VA}})$	(6) R&D Ind. (baseline)	(7)	(8) R&D Ind. (liberal)
human capital index	3.705**	3.613**	2.910**	4.185**	3.836***	0.213*	0.189**	0.240**
	(1.547)	(1.383)	(1.364)	(1.706)	(0.818)	(0.110)	(0.092)	(0.100)
ln(GDP per capita)	-0.381	-0.665*	-0.881***	-0.770**	-0.718	0.102***	0.090***	0.085***
	(0.389)	(0.369)	(0.319)	(0.355)	(0.447)	(0.028)	(0.027)	(0.028)
IPR protection	0.309	0.393**	0.383**	0.661***	0.027	0.029	0.024	0.028*
	(0.234)	(0.182)	(0.148)	(0.209)	(0.189)	(0.022)	(0.017)	(0.015)
R&D subsidies	0.695	0.577	0.380	0.261	0.720*	0.036	0.016	0.012
	(0.451)	(0.434)	(0.424)	(0.537)	(0.385)	(0.036)	(0.030)	(0.027)
ln (researchers)	0.249	0.416**	0.402**	0.214	0.217	0.071***	0.064***	0.073***
	(0.182)	(0.176)	(0.162)	(0.181)	(0.290)	(0.022)	(0.019)	(0.020)
log (sales)						0.006***	0.003**	0.004**
						(0.002)	(0.001)	(0.002)
Observations	7533	9672	11464	11464	7585	41396	71226	80253
R^2	0.668	0.679	0.688	0.703	0.653	0.618	0.562	0.603
Within R^2	0.014	0.016	0.013	0.016	0.019	0.008	0.006	0.005
Firm-period FE	Y	Y	Y	Y	Y	Y	Y	Y
Affiliate FE	Y	Y	Y	Y	Y	Y	Y	Y
Sample	mfg. only	excl. HQ	baseline	baseline	baseline	mfg. only	excl. HQ	baseline

Note: This table reports robustness results for the last two columns of Table A.8. The dependent variable in Columns 1-5 are the intensive margin invention intensity at an affiliate. Columns 1-2 use the baseline measure, but restrict the sample. Columns 3-5 use different measures for either invention or sales in measuring invention intensity. The dependent variable in Columns 6-8 is the indicator for R&D centers. Columns 6-7 use the baseline measure and restrict the sample; Column 8 uses the liberal definition of R&D centers. Standard errors (in parenthesis) are clustered two way, by firm and by host country. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table SA.A.2: Fact 2 Robustness: Manufacturing Only

Dependent variable	(1) prod. indicator	(2)	(3) log (sales)	(4)
R&D Indicator$_{fh,t}$	0.295***	1.042***		
	(0.004)	(0.033)		
ln(patent)$_{fh,t}$			0.324***	0.181***
			(0.014)	(0.046)
ln (distance)$_{fh,t}$		-0.012	-0.395**	
		(0.035)	(0.169)	
common language$_{fh,t}$		0.207***	0.284	
		(0.077)	(0.328)	
contiguity$_{fh,t}$		0.201***	0.227	
		(0.077)	(0.285)	
colonial tie$_{fh,t}$		-0.045	-0.713*	
		(0.070)	(0.382)	
Observations	4156173	61474	13072	6417
R^2	0.735	0.512	0.574	0.969
Within R^2	0.050	0.052	0.099	0.020
Firm-period FE	Y	Y	Y	Y
Host-period FE	Y	Y	Y	-
Home-host FE	Y	Y	Y	-
Host-industry FE	Y	-	-	-
Host-industry-period FE	-	-	-	Y
Affiliate FE	-	-	-	Y

Note: This table reports robustness of Fact 2 using only manufacturing firms. Column 1 corresponds to Column 1 of Table A.9; Columns 2 through 4 correspond to Columns 4, 5 and 7 of Table A.9. Standard errors (in parenthesis) are clustered by firm. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table SA.A.3: Fact 2 Robustness: Alternative Measures for Invention and Production

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable:	prod. indicator			ln(sales)				ln(value added)	
R&D Ind.$_{fh,t}$: liberal	0.275***	0.988***							
	(0.003)	(0.026)							
ln(citation)$_{fh,t}$			0.277***	0.137***					
			(0.012)	(0.040)					
ln(unwgt. patent)$_{fh,t}$					0.382***	0.194***			
					(0.015)	(0.046)			
R&D Ind.$_{fh,t}$: baseline							1.036***		
							(0.029)		
ln(patent)$_{fh,t}$								0.338***	0.107*
								(0.014)	(0.058)
ln (distance)$_{fh,t}$		-0.049*	-0.333**		-0.305**		-0.061*	-0.600***	
		(0.025)	(0.146)		(0.145)		(0.037)	(0.229)	
common language$_{fh,t}$		0.167***	0.513*		0.358		0.234***	0.827**	
		(0.051)	(0.272)		(0.268)		(0.061)	(0.330)	
contiguity$_{fh,t}$		0.096*	0.310		0.187		0.109**	-0.239	
		(0.049)	(0.240)		(0.237)		(0.054)	(0.258)	
colonial tie$_{fh,t}$		0.017	-0.657**		-0.552*		-0.051	-0.554	
		(0.046)	(0.314)		(0.307)		(0.056)	(0.363)	
Observations	7494979	119503	19519	8839	19519	8839	70184	12342	5644
R^2	0.705	0.495	0.559	0.963	0.568	0.963	0.554	0.601	0.930
Within R^2	0.048	0.045	0.066	0.015	0.086	0.020	0.065	0.132	0.005
Firm-period FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Host-period FE	Y	Y	Y	-	Y	-	Y	Y	-
Home-host FE	Y	Y	Y	-	-	-	Y	Y	-
Host-industry FE	Y	Y	Y	-	-	-	Y	Y	-
Host-industry-period FE	-	-	-	Y	-	Y	-	-	Y
Affiliate FE	-	-	-	Y	-	Y	-	-	Y

Note: This table reports robustness results for Fact 2 using alternative measures of invention and production. Columns 1 and 2 reproduce Columns 1 and 4 of Table A.9 using the liberal definition of R&D center; Column 7 reproduces Column 4 of Table A.9 using log value added as a measure for production. Columns 3 to 6 reproduces Columns 5 and 7 of Table A.9 using alternative intensive measures of invention. Columns 8 and 9 reproduce the same two columns using log value added to measure production. Standard errors (in parenthesis) are clustered by firm. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table SA.A.4: Fact 3 Robustness: Manufacturing Only

	(1)	(2)	(3)	(4)
	Affiliate R&D		Affiliate Production	
	indicator	ln(patent)	indicator	ln (sales)
ln(distance)$_{oh}$	-0.001	-0.140***	-0.004***	-0.230***
	(0.001)	(0.040)	(0.002)	(0.028)
common language$_{oh}$	0.022***	0.226**	0.011	0.005
	(0.006)	(0.093)	(0.009)	(0.066)
contiguity$_{oh}$	0.001	0.048	-0.001	0.137**
	(0.002)	(0.101)	(0.004)	(0.066)
colonial tie	0.006	0.094	0.029***	0.102
	(0.005)	(0.080)	(0.009)	(0.068)
R&D indicator$_{fh,t}$			0.379***	1.181***
			(0.022)	(0.037)
Observations	4045403	28244	4045403	54208
R^2	0.149	0.297	0.368	0.465
Within R^2	0.004	0.010	0.069	0.063
Firm-period FE	Y	Y	Y	Y
Host-industry FE	Y	Y	Y	Y
Host-period FE	Y	Y	Y	Y

Note: This table reproduces Columns 1-2 and 5-6 of Table A.10 for manufacturing firms. Measures of R&D and production are the same as in the baseline. Standard errors (in parenthesis) are clustered by country pair. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table SA.A.5: Fact 3: Alternative Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Headquarter Effect for R&D			Headquarter Effect for Production			
Dependent var.	R&D ind. (liberal)	ln(unwgt. patent)	ln(citation)	ln(sales)			ln(VA)
ln(distance)$_{oh}$	-0.003***	-0.119***	-0.078**	-0.247***	-0.102**	-0.131**	-0.189***
	(0.001)	(0.028)	(0.031)	(0.027)	(0.047)	(0.051)	(0.045)
common language$_{oh}$	0.030***	0.222***	0.224***	0.086	0.059	0.076	0.025
	(0.006)	(0.063)	(0.070)	(0.060)	(0.082)	(0.084)	(0.070)
contiguity$_{oh}$	0.005*	0.109*	0.183***	0.175***	-0.005	0.009	0.208***
	(0.002)	(0.062)	(0.065)	(0.057)	(0.093)	(0.098)	(0.066)
colonial tie$_{oh}$	0.001	0.025	-0.015	0.123*	0.046	0.060	0.156**
	(0.005)	(0.055)	(0.067)	(0.068)	(0.079)	(0.083)	(0.074)
R&D ind. (liberal)$_{fh,t}$				1.096***			
				(0.028)			
ln(unwgt. patent)$_{fh,t}$					0.392***		
					(0.018)		
ln(citation)$_{fh,t}$						0.284***	
						(0.015)	
R&D ind. (baseline)$_{fh,t}$							1.185***
							(0.030)
Observations	7295102	45364	45364	103131	16189	16189	60553
R^2	0.155	0.324	0.455	0.445	0.502	0.490	0.504
Within R^2	0.006	0.016	0.010	0.054	0.090	0.069	0.079
Firm-period FE	Y	Y	Y	Y	Y	Y	Y
Host-industry FE	Y	Y	Y	Y	Y	Y	Y
Host-period FE	Y	Y	Y	Y	Y	Y	Y

Note: This table reports the robustness results of Fact 3 to alternative measures. Columns 1 to 3 replicate Columns 1 and 2 of Table A.10 with different measures of R&D. Columns 4 to 6 replicate Column 6 of Table A.10 with different measures of R&D. Column 7 replicate Column 6 of Table A.10 using value added as a measure for production. Standard errors (clustered at country-pair level) in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

SA.A.3 IV Estimates for Fact 2

The second fact, the colocation of invention and production, plays an important role in quantitative analysis as it pins down the market access motive. The most demanding specification for this fact controls for firm-period, host-industry-period, and affiliate fixed effects, which rules out the following confounding factors: shocks to host countries or all affiliates of a firm that drive the colocation; idiosyncratic match quality between a firm and a host, which encourages both invention and production; changes in the comparative advantage of a country that affect the entire industry. A remaining source of threat is time changes in the idiosyncratic match quality between firms and hosts. Such changes need to be firm specific—otherwise it will be absorbed by host-industry-period fixed effects. One example of such shocks is development of a new technology in a country that is useful to only a few firms within an industry, but affects both production and invention of these firms directly.

To address this concern, I use an alternative identification strategy. Under the assumption that controlling for other time-varying country characteristics, changes in a host's R&D environment affect affiliate production only through affiliate R&D, proxies of R&D environment can serve as as instrumental variables. I use three instruments: R&D subsidies, the IPR protection index, and the number of researchers in the country. The first two are policies set at national level without regard to individual foreign firms; the last one depends largely on the supply factors. These variables might be correlated with other country-level determinants of affiliate production, which motivates me to control for host size, GDP per capita, and the general human capital measure. Finally, I also include affiliate and other fixed effects, so identification comes only from time variation within a host.

Columns 1 and 2 of Table SA.A.6 report the baseline 2SLS results and the corresponding first stage regression. The first stage shows that R&D subsidies, the number of researchers, and IPR protection all have positive effects on affiliate R&D. The robust F statistic is above 10, the conventional rule of thumb for detecting weak IV. The 2SLS estimate suggests that a one-percent increase in affiliate R&D increases production by 0.46%, which is similar to the baseline estimate.

The lower panel of the table reports additional diagnostic statistics. Recent studies (c.f. Lee, McCrary, Moreira and Porter, 2020) suggest that 2SLS inference based on standard t-statistics might not be conservative enough. I report tests from two tests that are robust to weak IV, both of which are able to reject that the coefficient is zero. Finally, the over-identification test suggests that we cannot reject that all three IVs give the same estimate.

One might be concerned that IPR protection can affect the sales of affiliates directly—in host counties experiencing an improvement in the protection of IPR, the risk of IP theft is lower, so firms might be more willing to move production there. Columns 3 and 4 allow IPR protection to be endogenous and use the remaining two IVs. The results are similar. The first-stage is slightly weaker, but the weak-IV robust tests are both able to reject that the key coefficient is zero.

I conduct additional robustness exercises: Columns 5 and 6 focus on manufacturing firms; Columns 7 through 10 use two alternative measures of invention—unweighted patent counts and citations; Columns 11 and 12 use value added, instead of sales, to measure firm production. All these robustness exercises give similar results.

This IV strategy has its weaknesses: despite other time-varying controls, it is still possible that R&D subsidies and the number of researchers in a country are correlated with unobserved changes in country characteristics that directly affect the sales of affiliates. But to the extent that this is the main concern, it is directly addressed in the baseline specifications with host-industry-period fixed effects. In this sense, while both identification strategies are imperfect, they exploit orthogonal variations and thus complement each other. That both approaches give qualitatively similar estimates is reassuring.

Table SA.A.6: Fact 2: IV Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Baseline IV		Endogenous IPR		Mfg. Only		R&D Measure 2		R&D Measure 3		Prod. Measure 2	
	2SLS	1st stage	2SLS	1st stage	2SLS	1st Stage	2SLS	1st stage	2SLS	1st stage	2SLS	1st stage
ln(patent)$_{fh,t}$	0.466***		0.516**		0.599***						0.673*	
	(0.141)		(0.240)		(0.160)						(0.354)	
ln(unwgt. patent)$_{fh,t}$							0.533***					
							(0.166)					
ln(citation)$_{fh,t}$									0.451**			
									(0.183)			
R&D subsidies		0.374		0.374		0.436*		0.181		0.060		0.304
		(0.228)		(0.228)		(0.265)		(0.210)		(0.329)		(0.186)
ln (researchers)		0.635***		0.635***		0.514***		0.607***		0.409**		0.562***
		(0.170)		(0.170)		(0.182)		(0.137)		(0.202)		(0.170)
IPR protection		0.459**	-0.062	0.459**		0.441**		0.431***		0.703***		0.139
		(0.191)	(0.231)	(0.191)		(0.220)		(0.137)		(0.187)		(0.169)
Observations	11464	11464	11464	11464	7533	7533	11464	11464	11464	11464	7585	7585
Time-varying controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Firm-period FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Home-Host FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Affiliate FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
1st Stage F												
K-P F-stat		10.380		7.448		8.216		14.703		12.324		7.304
Weak IV robustness inference												
Anderson-Rubin test p-value	$<$0.01		0.028		$<$0.01		$<$0.01		$<$0.01		$<$0.01	
Stock-Wright test p-value	0.025		0.086		0.036		0.025		0.025		0.072	
Over identification test												
Hansen J statistic p-value	0.461		0.227		0.389		0.565		0.299		0.266	

Note: This reports the estimates for colocation between invention and production, using time variation in host R&D subsidies, the number of researchers, and the IPR protection index as instrumental variables for changes in affiliate invention. All specification includes firm-period, home-home, and affiliate fixed effects and the following time varying controls: ln(GDP), ln(GDP per capita), and the human capital index. Columns 1 and 2 use all three IVs and the baseline sample; Columns 3 and 4 allows the IPR protection index to be endogenous; Columns 5 and 6 restrict to manufacturing industry only; Column 7 to 10 use two alternative measures of affiliate invention: unweighted patent counts and citation; Columns 11 and 12 use value added to measure affiliate production as the outcome variable. Standard errors (in parenthesis) are clustered two way, by host country and by firm. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

SA.A.4 Suggestive Evidence on the Independence among R&D Centers

The quantitative model developed in Section 3 assumes that offshore R&D centers belonging to the same parent develop differentiated varieties and thus operate independently. This subsection provides evidence in support of this assumption.

Specifically, I investigate whether the R&D decision of an MNC in host i responds to changes in R&D-related policies or other factors in the headquarters o and in other countries $i' \neq i$ where the firm has a presence. The idea is as follows: Section SA.A.3 of this appendix shows that R&D in host i responds to R&D-related shocks in country i. If the coordination among R&D centers is important, in response to the expansion of R&D in i, firms will adjust R&D in other hosts.[3]

Table SA.A.7 reports the results. Columns 1 and 2 regress the extensive and intensive margin measures of offshore R&D on the characteristics of the headquarter country, with headquarters themselves excluded from the sample. I control for affiliate and host-period fixed effects, so the coefficients are identified off time variation among affiliates from different countries. None of the predictors of R&D—R&D subsidies, IPR protection, the number of researchers—has a statistically significant impact on affiliate R&D. This statistical insignificance is not due to the lack of variation in these characteristics: in fact, as the first-stage regressions reported in Table SA.A.6 shows, the same variables have economically sizeable and statistically significant impacts on R&D in the same host.

Columns 3 to 6 regress affiliate R&D on the average characteristics of all other countries in which the

[3] I focus on the response to shocks, rather than on the cross-sectional relation, because the model implies that the affiliates of the same parent inherit correlated innovation efficiency, hence their invention output might be correlated even if each of them carry out R&D independently.

firm has a presence. Because firms operate in different sets of countries, the variation is at firm level and we can control for home-period fixed effects. Columns 3 and 4 include headquarters in the regression sample; Columns 5 and 6 exclude headquarters. Both set of regressions find that affiliate R&D do not respond to shocks affecting firm's R&D in either the home or other host countries.

Although I can not rule out that firms coordinate among their affiliates on R&D entirely, this piece of evidence suggests that such coordination is not a first-order feature of my data.

Table SA.A.7: Evidence on R&D Center Independence

	(1)	(2)	(3)	(4)	(5)	(6)
	R&D and HQ shocks		R&D and shocks in other countries			
			incl. HQ		excl. HQ	
	R&D ind.	ln(patent)	R&D ind.	ln(patent)	R&D ind.	ln(patent)
Headquarter country char.:						
R&D subsidies	-0.000	0.122				
	(0.001)	(0.144)				
IPR protection	-0.001	0.007				
	(0.001)	(0.130)				
ln (researcher)	-0.000	0.006				
	(0.001)	(0.118)				
Average char. of other countries where firm has an affiliate:						
R&D subsidies			-0.039	-0.109	-0.072	-0.145
			(0.037)	(0.182)	(0.046)	(0.390)
IPR protection			0.007	0.008	0.012	0.111
			(0.009)	(0.041)	(0.011)	(0.086)
ln (researcher)			0.002	0.009	0.002	0.006
			(0.001)	(0.006)	(0.001)	(0.010)
Observations	3290430	21216	112731	34012	97773	21966
R^2	0.571	0.740	0.724	0.768	0.666	0.707
Within R^2	0.000	0.001	0.000	0.003	0.000	0.003
Time-varying controls	Y	Y	Y	Y	Y	Y
Affiliate FE	Y	Y	Y	Y	Y	Y
Host-period FE	Y	Y	Y	Y	Y	Y
Home-period FE	-	-	Y	Y	Y	Y

Note: Columns 1 and 2 regress measures of affiliate R&D on time-varying characteristics of the home country. Additional time-varying controls are ln(GDP), ln(GDP per capita), and the human capital index of the home country. Headquarters themselves are excluded from the regression. Columns 3 to 6 regress measures of affiliate R&D on the average characteristics among all other countries in which the firm has an affiliate. Additional time-varying controls are the average value of ln(GDP), ln(GDP per capita), and the human capital index of these countries. Columns 3 and 4 include headquarters as one of the affiliates; Columns 5 and 6 exclude the headquarters themselves. Affiliate fixed effects are controlled for throughout. Standard errors in parenthesis. Standard errors for Columns 1 and 2 are clustered two way, by home country and by firm; standard errors in Columns 3-5 are clustered by firm. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

SA.B Theory

SA.B.1 Proof of Lemma B.1 in the Appendix

Proof. 1. $\forall x$, following the definition of ζ

$$
\begin{aligned}
Prob(\zeta \leq x|z) &= Pr(A_1\eta_1 \leq x, ..., A_N\eta_N \leq x|z) \\
&= Prob(\eta_1 \leq \frac{x}{A_1}, ..., \eta_N \leq \frac{x}{A_N}|z) \\
&\text{(using the definition of } H(\cdot|z)) \\
&= \begin{cases} 1 - \Big(\sum_m \frac{z^\theta}{N}[(\frac{x}{A_m})^{-\theta}]\Big) = 1 - \tilde{A}^\theta x^{-\theta}, \text{ if } x \geq z \cdot \max_m A_m \equiv \bar{A} \\ 0, \quad \text{if } x < \bar{A}. \end{cases}
\end{aligned}
$$

2. From equation (B.3), for $x \geq \bar{A}$, from part 1, $\forall x' \geq x$

$$
Pr(\zeta > x'|\zeta > x) = \frac{(\frac{x'}{\bar{A}})^{-\theta}}{(\frac{x}{\bar{A}})^{-\theta}} = \Big(\frac{x'}{x}\Big)^{-\theta}.
$$

Therefore, the conditional distribution of ζ above $\forall x \geq \bar{A}$ is Pareto with tail parameter θ and scale parameter x. Thus we have

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

3. For $x > \bar{A}$,

$$
Pr(m = \arg\max_{m'} A_{m'}\eta_{m'} \wedge A_m\eta_m \geq x|z) = \int_x^\infty Prob(A_{m'}\eta_{h'} \leq u, \forall m' \neq m|A_m\eta_m = u, z)f_m(u|z)du,
$$

where $f_m(u|z)$ is the marginal density of $A_m\eta_m$ conditional on z. For $u \geq x > \bar{A}$, the integrand in the above function is:

$$
\begin{aligned}
Prob(A_{m'}\eta_{m'} \leq u, \forall m' \neq m|A_m\eta_m = u, z)f_m(u|z) &= \frac{\partial Pr(A_1\eta_1 \leq u, A_m\eta_m \leq C, ...A_N\eta_N \leq u|z)}{\partial C}\Big|_{C=u} \\
&= z^\theta \frac{A_m^\theta}{N}\theta u^{-\theta-1}.
\end{aligned}
$$

Therefore,

$$
Pr(m = \arg\max_{m'} A_{m'}\eta_{m'} \wedge \zeta \geq x|z) = z^\theta \frac{A_m^\theta}{N}x^{-\theta}.
$$

And

$$
\begin{aligned}
Pr(m = \arg\max_{m'} A_{m'}\eta_{m'} \Big| \zeta \geq x, z) &= \frac{Pr(m = \arg\max_{m'} A_{m'}\eta_{m'} \wedge \zeta \geq x|z)}{Pr(\zeta \geq x|z)} \\
&= \frac{z^{\theta}\frac{A_m^{\theta}}{N}x^{-\theta}}{\left(z^{\theta}\frac{1}{N}\sum_{m'} A_{m'}^{\theta}\right)x^{-\theta}} \\
&= \frac{A_m^{\theta}}{\sum_{m'} A_{m'}^{\theta}},
\end{aligned}
$$

Note that $\forall x' \geq x > \bar{A}$

$$
\begin{aligned}
Prob(\zeta \geq x'|m = \arg\max_{m'} A_{m'}\eta_{m'} \wedge \zeta \geq x, z\Big) &= \frac{Pr(m = \arg\max_{m'} A_{m'}\eta_{m'} \wedge \zeta \geq x' \wedge \zeta \geq x|z)}{Pr(m = \arg\max_{m'} A_{m'}\eta_{m'} \wedge \zeta \geq x|z)} \\
&= \frac{z^{\theta}\frac{A_m^{\theta}}{N}x'^{-\theta}}{z^{\theta}\frac{A_m^{\theta}}{N}x^{-\theta}} \\
&= (\frac{x'}{x})^{-\theta}.
\end{aligned}
$$

□

SA.B.2 Proof of Proposition 1

Proof. 1. From equations (3) and (6), we can write the expected *per-variety* profit for a firm from o with efficiency z^P as

$$
\overline{\pi}_{oi}(z^P) \propto (z^P)^{\theta} \cdot \sum_d \underbrace{X_d^{\frac{\theta}{\sigma-1}} P_d^{\theta} (W_d^h f_d^M)^{\frac{\theta+1-\sigma}{1-\sigma}}}_{\equiv \tilde{X}_d} (\tilde{\zeta}_{oid})^{\theta} \tag{SA.B.1}
$$

$$
\equiv (z^P)^{\theta} \cdot \sum_d \tilde{X}_d (\tilde{\zeta}_{oid})^{\theta}, \text{ where } \tilde{\zeta}_{oid} = [\sum_m \frac{1}{N} (\frac{T_m \cdot \phi_{oim}^p}{W_m^l \tau_{md}})^{\theta}]^{\frac{1}{\theta}}.
$$

$\phi_{oim}^P = \tau_{md} = 1 \; \forall o, i, m, d \implies \tilde{\zeta}_{oid}$ is a common across $o, i, d \implies$ we can write $\overline{\pi}_{oi}(z^P) \equiv \overline{\pi} \cdot (z^P)^{\theta}$, with $\overline{\pi}$ common to all host countries. So conditioning on z^P, the profit from a variety is the same regardless of where the firm is from or where R&D takes place.

We prove by contradiction that $W_i^h > W_{i'}^h$ and $L_i^h > L_{i'}^h$. Suppose first $W_i^h \leq W_{i'}^h$, then $L_i^h > L_{i'}^h$ for two reasons. First, as the profit from each variety is the same and foreign firms can more easily transfer knowhow to i than to i', if $W_i^h \leq W_{i'}^h$, more firms will enter i than i'. Second, all firms active in both i and i' face lower wage and have higher knowhow in i, so they recruit weakly more researchers. Skilled labor market clearing implies $L_i^h > L_{i'}^h$, which in turn implies that $\hat{\alpha}_i < \hat{\alpha}_{i'}$, $W_i^l = W_i^h \hat{\alpha}_i < W_{i'}^h \hat{\alpha}_{i'} = W_{i'}^l$. Since trade is frictionless, and i and i' have the same manufacturing productivity, the demand for manufacturing labor in i is higher than that in i'. Unskilled labor market clearing implies $L_i^l > L_{i'}^l$, which contradicts $L_i^h > L_{i'}^h$.

Therefore we must have $W_i^h > W_{i'}^h$. Suppose $L_i^h \leq L_{i'}^h$, then $\hat{\alpha}_i > \hat{\alpha}_{i'}$, which implies $W_i^l > W_{i'}^l$. Trade and manufacturing labor market clearing in turn implies $L_i^l < L_{i'}^l$, contradicting $L_i^h \leq L_{i'}^h$.

It follows that $W_i^h > W_{i'}^h$ and $L_i^h > L_{i'}^h$ holds. This means $\hat{\alpha}_i < \hat{\alpha}_{i'}$ and $L_i^l < L_{i'}^l$.

$$
\frac{W_i^h L_i^h}{W_i^l L_i^L} = \frac{L_i^h}{\hat{\alpha}_i L_i^l} > \frac{L_{i'}^h}{\hat{\alpha}_{i'} L_{i'}^l} = \frac{W_{i'}^h L_{i'}^h}{W_{i'}^l L_{i'}^L},
$$

so country i is more specialized in R&D relative to country i'.

Now we prove that domestic firms employ a smaller share of researchers in i than in i'. Using equations (2) and (4), the total demand for researchers in i is

$$L_i^h = \sum_o R_{oi} \int_{\hat{z}_{oi}}^{\infty} [\int_0^{\infty} \left(\frac{\gamma \overline{\pi}_{oi}(z^P) z^R}{W_i^h}\right)^{\frac{1}{1-\gamma}} dG^P(z^P|z^R)] dG_{oi}^R(z^R) + \sum_{o \neq i} R_{oi} f_{oi}^R \tag{SA.B.2}$$

$$= \left(\frac{\gamma \overline{\pi}}{W_i^h}\right)^{\frac{1}{1-\gamma}} \sum_o E_o (\phi_{oi}^R)^{1-\gamma} \int_{\hat{z}_{oi}}^{\infty} (\tilde{z}^R)^{\frac{1}{1-\gamma}} [\int_0^{\infty} (z^P)^{\frac{\theta}{1-\gamma}} dG^P(z^P|\tilde{z}^R \phi_{oi}^R)] dG_o^E(\tilde{z}^R) + \sum_{o \neq i} E_o \cdot \left(1 - G_o^E(\hat{z}_{oi}^R)\right) \cdot f_{oi}^R.$$

So the ratio of researchers employed at foreign affiliates over domestic firms is:

$$\frac{\sum_{o \neq i} E_o (\phi_{oi}^R)^{1-\gamma} \int_{\hat{z}_{oi}}^{\infty} (\tilde{z}^R)^{\frac{1}{1-\gamma}} [\int_{\mathbb{Z}^P} (z^P)^{\frac{\theta}{1-\gamma}} dG^P(z^P|\tilde{z}^R \phi_{oi}^R)] dG_o^E(\tilde{z}^R) + \sum_{o \neq i} E_o \cdot \left(1 - G_o^E(\hat{z}_{oi}^R)\right) \cdot f_{oi}^R \cdot (\frac{W_i^h}{\gamma \overline{\pi}})^{1-\gamma}}{E_i \int_0^{\infty} (\tilde{z}^R)^{\frac{1}{1-\gamma}} [\int_0^{\infty} (z^P)^{\frac{\theta}{1-\gamma}} dG^P(z^P|\tilde{z}^R)] dG_o^E(\tilde{z}^R)}.$$

The denominator of the ratio is the same in i and i'. The numerator of the ratio is higher in i as more firms enter (more researcher employed for the fixed entry cost) and each foreign entrant possess more knowhow in i. Thus, the ratio is higher in i than in i'.

2. As headquarters do not matter for production and trade is frictionless, $\tilde{\zeta}_{oid}$ depends only on i. Following the same argument as equation (SA.B.1), we can write $\overline{\pi}_{oi}$ as $\overline{\pi}_i$ for short.

   I first prove by contradiction that $\frac{\overline{\pi}_i}{W_i^h} > \frac{\overline{\pi}_{i'}}{W_{i'}^h}$. Suppose instead $\frac{\overline{\pi}_i}{W_i^h} \leq \frac{\overline{\pi}_{i'}}{W_{i'}^h}$.

   From equations (2) and (3), $h_{oi}(z^P, z^R) \propto (z^R)^{\frac{1}{1-\gamma}} \cdot (z^P)^{\frac{\theta}{1-\gamma}} \cdot (\frac{\overline{\pi}_i}{W_i^h})^{\frac{1}{1-\gamma}}$, we have that $\forall o, h_{oi}(z^P, z^R) < h_{oi'}(z^P, z^R)$. As i and i are otherwise identical, a firm operating R&D centers in both i and i' with have the same z^R in both countries. Conditional on their realization of z^P also being the same, the firm will hire more researcher in i' and have more varieties in i'.

   Now consider the entry into these two countries, governed by $\hat{z}_{oi}^R$ and $\hat{z}_{oi'}^R$

$$f_{oi}^R \propto \left(\frac{\overline{\pi}_i}{W_i^h}\right)^{\frac{1}{1-\gamma}} \cdot (\hat{z}_{oi}^R \phi_{oi}^R)^{\frac{1}{1-\gamma}} \cdot \int_0^{\infty} (z^P)^{\frac{\theta}{1-\gamma}} \cdot dG^P(z^P|\hat{z}_{oi}^R \phi_{oi}^R)$$

$$f_{oi'}^R \propto \left(\frac{\overline{\pi}_{i'}}{W_{i'}^h}\right)^{\frac{1}{1-\gamma}} \cdot (\hat{z}_{oi'}^R \phi_{oi'}^R)^{\frac{1}{1-\gamma}} \cdot \int_0^{\infty} (z^P)^{\frac{\theta}{1-\gamma}} \cdot dG^P(z^P|\hat{z}_{oi'}^R \phi_{oi'}^R)$$

   Noting that $\forall o,\ f_{oi}^R = f_{oi'}^R, \phi_{oi}^R = \phi_{oi'}^R$ and that the integral increases in ϕ_{oi}^R, we have

$$\frac{\overline{\pi}_i}{W_i^h} \leq \frac{\overline{\pi}_{i'}}{W_{i'}^h} \implies \hat{z}_{oi}^R > \hat{z}_{oi'}^R, \forall o \neq i, i'$$

   i.e., more firms enter country i' than country i for R&D.

   Since more firms enter i' for R&D and firms active in both recruit more researchers in i', the total demand for researchers is higher in i', so we have $L_i^h < L_{i'}^h$, $\hat{\alpha}_i > \hat{\alpha}_{i'}$. This in turn implies $L_i^l > L_{i'}^l$.

   Now consider the total low-skill value added in m. Since trade is frictionless, the total manufacturing production in a country m is

$$
\begin{aligned}
Y_m &\propto \sum_o \sum_{\tilde{i}} \sum_d (W_d^h f_d^M)^{\frac{\theta-(\sigma-1)}{1-\sigma}} \left(\frac{X_d}{P_d^{1-\sigma}}\right)^{\frac{\theta}{\sigma-1}} \left(\frac{T_m \phi^P_{o\tilde{i}m}}{W^l_m}\right)^\theta \int_0^\infty (z^P)^\theta V_{o\tilde{i}}(z^P) dz^P. \\
&= \Big[\sum_d (W_d^h f_d^M)^{\frac{\theta-(\sigma-1)}{1-\sigma}} \left(\frac{X_d}{P_d^{1-\sigma}}\right)^{\frac{\theta}{\sigma-1}}\Big] \sum_o \sum_{\tilde{i}} \left(\frac{T_m \phi^P_{o\tilde{i}m}}{W^l_m}\right)^\theta \int_0^\infty (z^P)^\theta V_{o\tilde{i}}(z^P) dz^P.
\end{aligned}
$$

Suppose $W_i^l \geq W_{i'}^l$

$$
\begin{aligned}
&Y_{i'} - Y_i \propto \sum_o \sum_{\tilde{i}} \underbrace{\left(\int_0^\infty (z^P)^\theta V_{o\tilde{i}}(z^P) dz^P\right)}_{\equiv V_{o\tilde{i}}} \Big[\left(\frac{T_{i'}\phi^P_{o\tilde{i}i'}}{W^l_{i'}}\right)^\theta - \left(\frac{T_i \phi^P_{o\tilde{i}i}}{W^l_i}\right)^\theta\Big] \qquad \text{(SA.B.3)} \\
&= \sum_o \sum_{\tilde{i}\neq i,i'} V_{o\tilde{i}} \Big[\left(\frac{T_{i'}\phi^P_{o\tilde{i}i'}}{W^l_{i'}}\right)^\theta - \left(\frac{T_i \phi^P_{o\tilde{i}i}}{W^l_i}\right)\Big] + \sum_o \Big[V_{oi}\left(\frac{T_{i'}\phi^P_{oii'}}{W^l_{i'}}\right)^\theta + V_{oi'}\left(\frac{T_{i'}\phi^P_{oi'i'}}{W^l_{i'}}\right)^\theta - V_{oi}\left(\frac{T_i\phi^P_{oii}}{W^l_i}\right)^\theta - V_{oi'}\left(\frac{T_i\phi^P_{oi'i}}{W^l_i}\right)^\theta\Big] \\
&> \sum_o \Big[V_{oi}\left(\frac{T_{i'}\phi^P_{oii'}}{W^l_{i'}}\right)^\theta + V_{oi'}\left(\frac{T_{i'}\phi^P_{oi'i'}}{W^l_{i'}}\right)^\theta - V_{oi}\left(\frac{T_i\phi^P_{oii}}{W^l_i}\right)^\theta - V_{oi'}\left(\frac{T_i\phi^P_{oi'i}}{W^l_i}\right)^\theta\Big] \\
&= \sum_o \Big[V_{oi} \underbrace{\left(\left(\frac{T^P_{i'}\phi^P_{oi'i'}}{W^l_{i'}}\right)^\theta - \left(\frac{T^P_i\phi^P_{oii}}{W^l_i}\right)^\theta\right)}_{>0} + V_{oi} \underbrace{\left(\left(\frac{T^P_{i'}\phi^P_{oii'}}{W^l_{i'}}\right)^\theta - \left(\frac{T^P_i\phi^P_{oi'i}}{W^l_i}\right)^\theta\right)}_{>0} + (V_{oi'} - V_{oi}) \underbrace{\left(\left(\frac{T^P_{i'}\phi^P_{oi'i'}}{W^l_{i'}}\right)^\theta - \left(\frac{T^P_i\phi^P_{oi'i}}{W^l_i}\right)^\theta\right)}_{>0}\Big] \\
&> 0,
\end{aligned}
$$

where the last inequality uses that $T_i = T_{i'}$, $V_{oi'} > V_{oi}$, $\phi^P_{oi'i'} = \phi^P_{oii} = 1$, and $1 > \phi^P_{oii'} > \phi^P_{oi'i} > 0$.

Thus we have proved $Y_{i'} > Y_i$. Together with $W^l_{i'} \leq W^l_i$, this implies $L^l_{i'} > L^l_i$, which leads to contradiction.

Now consider $W^l_i \leq W^l_{i'}$. Since $\hat{\alpha}_i > \hat{\alpha}_{i'}$, we have $W^h_i < W^h_{i'}$. Observe that

$$
\overline{\pi}_i \geq \overline{\pi}_{i'} \iff \sum_m \left(\frac{T_m \cdot \phi^p_{oim}}{W^l_m \tau_{md}}\right)^\theta \geq \sum_m \left(\frac{T_m \cdot \phi^p_{oi'm}}{W^l_m \tau_{md}}\right)^\theta
$$

.

$$
\begin{aligned}
\sum_m \left(\frac{T_m \cdot \phi^p_{oim}}{W^l_m \tau_{md}}\right)^\theta - \sum_m \left(\frac{T_m \cdot \phi^p_{oi'm}}{W^l_m \tau_{md}}\right)^\theta &= \sum_{m\neq i,i'} \Big[\left(\frac{T_m \cdot \phi^p_{oim}}{W^l_m \tau_{md}}\right)^\theta - \left(\frac{T_m \cdot \phi^p_{oi'm}}{W^l_m \tau_{md}}\right)^\theta\Big] \\
&+ \left(\frac{T_i \cdot \phi^p_{oii}}{W^l_i \tau_{id}}\right)^\theta + \left(\frac{T_{i'} \cdot \phi^p_{oii'}}{W^l_{i'} \tau_{i'd}}\right)^\theta - \left(\frac{T_i \cdot \phi^p_{oi'i}}{W^l_i \tau_{id}}\right)^\theta - \left(\frac{T_{i'} \cdot \phi^p_{oi'i'}}{W^l_{i'} \tau_{i'd}}\right)^\theta \\
&> 0,
\end{aligned}
$$

where the inequality uses that trade is frictionless, that i has better access than i' to other production locations $m \neq i, i'$, and that i and i' are otherwise identical.

This implies $\overline{\pi}_i \geq \overline{\pi}_{i'}$. Together with $W^h_i < W^h_{i'}$, it contradicts the premise that $\frac{\overline{\pi}_i}{W^h_i} \leq \frac{\overline{\pi}_{i'}}{W^h_{i'}}$.

Thus, we have proved that $\frac{\overline{\pi}_i}{W^h_i} > \frac{\overline{\pi}_{i'}}{W^h_{i'}}$. Per the analysis above, this also implies that more firms will enter i for R&D, and each of them will invest more in R&D. Because country i will have more foreign firms, a higher fraction of researchers will be at foreign affiliates.

3. We first prove by contradiction that $W_i^h \leq W_{i'}^h$. Suppose instead that $W_i^h > W_{i'}^h$. Since MP is frictionless, $\overline{\pi}_{oi}(z^P)$ is independent of both o and i, i.e., $\overline{\pi}_{oi} = \overline{\pi}_{oi'}, \forall o \implies \frac{\overline{\pi}_{oi}}{W_i^h} < \frac{\overline{\pi}_{oi'}}{W_{i'}^h}$.

   Following the same rationale as in the proof of Part 2 of this proposition, we have $L_i^h < L_{i'}^h$, $L_i^l > L_{i'}^l$, and $\hat{\alpha}_i > \hat{\alpha}_{i'}$.

   Because of the trade cost specification, if $W_i^l \geq W_{i'}^l$, any destination location will spend weekly less on the goods produced in i than the goods produced in i', which contradicts labor market clearing condition and $W_i^l L_i^l > W_{i'}^l L_{i'}^l$.

   Summing up, we have $W_i^l < W_{i'}^l$, $W_i^h > W_{i'}^h$, and $\hat{\alpha}_i > \hat{\alpha}_{i'}$, which contradicts with $W_i^h \hat{\alpha}_i = W_i^l$.

   It follows from this contradiction that $W_i^h < W_{i'}^h \implies \frac{\overline{\pi}_{oi}}{W_i^h} > \frac{\overline{\pi}_{oi'}}{W_{i'}^h}$. Following the same rationale as above, this means a higher fraction of workers in i will work in research, and that more foreign R&D centers will enter i. As firms' research expenditures are proportional to their knowhow (i.e., $\frac{h_{oi}(z^P,z^R)}{h_{oi}(\tilde{z}^P,\tilde{z}^R)} \propto (\frac{z^R}{\tilde{z}^R})^{\frac{1}{1-\gamma}} \cdot (\frac{z^P}{\tilde{z}^P})^{\frac{\theta}{1-\gamma}}$), this also mean a higher share of researchers will be at foreign firms.

   □

SA.B.3 Proof of Proposition 2

Proof. Following equation (SA.B.1), frictionless MP $\implies$ we can write $\overline{\pi}_{oi}(z^P) \equiv \overline{\pi} \cdot (z^P)^\theta$, with $\overline{\pi}$ being a constant. The demand for researchers in city i is given by equation (SA.B.2) as below:

$$L_i^h = \tag{SA.B.4}$$
$$\left(\frac{\gamma\overline{\pi}}{W_i^h}\right)^{\frac{1}{1-\gamma}} \sum_o E_o(\phi_{oi}^R)^{1-\gamma} \int_{\hat{z}_{oi}^R}^\infty (\tilde{z}^R)^{\frac{1}{1-\gamma}} [\int_0^\infty (z^P)^{\frac{\theta}{1-\gamma}} dG^P(z^P|\tilde{z}^R\phi_{oi}^R)] dG_o^E(\tilde{z}^R) + \sum_{o\neq i} E_o \cdot \left(1 - G_o^E(\hat{z}_{oi}^R)\right) \cdot f_{oi}^R.$$

1. We prove that $W_i^h < W_{i'}^h$ by contradiction. Suppose $W_i^h \geq W_{i'}^h$. The entry decision to a host country i is determined by

$$f_{oi}^R \propto \left(\frac{\overline{\pi}_{oi}}{W_i^h}\right)^{\frac{1}{1-\gamma}} \cdot (\hat{z}_{oi}^R \phi_{oi}^R)^{\frac{1}{1-\gamma}} \cdot \int_0^\infty (z^P)^{\frac{\theta}{1-\gamma}} \cdot dG^P(z^P|\hat{z}_{oi}^R\phi_{oi}^R)$$
$$f_{oi'}^R \propto \left(\frac{\overline{\pi}_{oi'}}{W_{i'}^h}\right)^{\frac{1}{1-\gamma}} \cdot (\hat{z}_{oi'}^R \phi_{oi'}^R)^{\frac{1}{1-\gamma}} \cdot \int_0^\infty (z^P)^{\frac{\theta}{1-\gamma}} \cdot dG^P(z^P|\hat{z}_{oi'}^R\phi_{oi'}^R)$$

   With $\overline{\pi}_{oi}(z^P) = \overline{\pi}_{oi'}(z^P) = \overline{\pi} \cdot (z^P)^\theta$, $W_i^h \geq W_{i'}^h$ implies $\forall o \neq i, i', \hat{z}_{oi}^R > \hat{z}_{oi'}^R$.

   $W_i^h \geq W_{i'}^h$ and $\hat{z}_{oi}^R > \hat{z}_{oi'}^R$ jointly imply $L_i^h < L_{i'}^h$. Under the assumption that $A_i(\alpha)$ first order stochastically dominates $A_{i'}(\alpha)$, $L_i^h < L_{i'}^h \implies \hat{\alpha}_i > \hat{\alpha}_{i'}$ and $L_i^l > L_{i'}^l$.

   In turn, we have $W_i^h \hat{\alpha}_i > W_{i'}^h \hat{\alpha}_{i'} \implies W_i^l > W_{i'}^l$. Since i and i' face the same export costs to all destinations and have the same manufacturing efficiency T, $W_i^l > W_{i'}^l \implies L_i^l < L_{i'}^l$, a contradiction.

   Thus, we have proved that $W_i^h < W_{i'}^h$. Following the same argument about the entry decision, we have $\forall o \neq i, i', \hat{z}_{oi}^R < \hat{z}_{oi'}^R$. This means that more firms will do R&D in i than in i'. As firms' research expenditures are proportional to their knowhow (i.e., $\frac{h_{oi}(z^P,z^R)}{h_{oi}(\tilde{z}^P,\tilde{z}^R)} \propto (\frac{z^R}{\tilde{z}^R})^{\frac{1}{1-\gamma}} \cdot (\frac{z^P}{\tilde{z}^P})^{\frac{\theta}{1-\gamma}}$), and because of the fixed cost for offshore R&D, more foreign firms in i also mean a higher share of researchers will be at foreign firms in i.

2. We start by showing that $W_i^h > W_{i'}^h$. Suppose instead $W_i^h \leq W_{i'}^h$. More foreign firms will enter i for R&D than i'. As i also has more domestic knowhow, from equation (SA.B.4), $L_i^h > L_{i'}^h$. Since the two countries have the same talent distribution, $L_i^h > L_{i'}^h \implies \hat{\alpha}_i < \hat{\alpha}_{i'}$ and $L_i^l < L_{i'}^l$.

This means $\hat{\alpha}_i W_i^h < W_{i'}^h \hat{\alpha}_{i'}$, i.e., $W_i^l < W_{i'}^l$. Since i and i' face the same export costs to all destinations and have the same manufacturing efficiency T, $W_i^l < W_{i'}^l$ implies that $L_i^l > L_{i'}^l$, a contradiction.

Thus we have proved that $W_i^h > W_{i'}^h$. Suppose that $\hat{\alpha}_i \geq \hat{\alpha}_{i'}$, then $W_i^l > W_{i'}^l$, which implies $L_i^l < L_{i'}^l$, contradicting $\hat{\alpha}_i \geq \hat{\alpha}_{i'}$. Therefore we must have $\hat{\alpha}_i < \hat{\alpha}_{i'}$, so country i specializes in innovation. $W_i^h > W_{i'}^h$ also means that there will be fewer foreign firms entering country i than country i' which, in turn, implies that a smaller share of R&D in i will be at foreign affiliates.

3. Under free trade and MP, country d's spending that goes to production in m is:

$$\lambda_{md} = \frac{(T_m/W_m^l)^\theta}{\sum_{m'}(T_{m'}/W_{m'}^l)^\theta} \equiv \lambda_m,$$

which is common across destination markets.

$$\tilde{\zeta}_{oid} \equiv [\sum_m \frac{1}{N}(\frac{T_m \phi_{oim}^p}{W_m^l \tau_{md}})^\theta]^{\frac{1}{\theta}} = [\sum_m \frac{1}{N}(\frac{T_m}{W_m^l})^\theta]^{\frac{1}{\theta}} \equiv \tilde{\zeta}.$$

Derive the relationship on occupation choice. Let $\hat{\alpha}_i$ be the cut off in country i, then we have

$$\hat{\alpha}_i = \frac{W_i^l}{W_i^h} \tag{SA.B.5}$$

$$\frac{L_i^h}{L_i} = \int_{\hat{\alpha}_d}^\infty \alpha \cdot dA_i(\alpha) = \frac{\kappa_\alpha}{1-\kappa_\alpha}(\underline{\alpha}_i)^{\kappa_\alpha}\int_{\hat{\alpha}_d}^\infty d\alpha^{1-\kappa_\alpha} = \frac{\kappa_\alpha}{\kappa_\alpha - 1}(\underline{\alpha}_i)^{\kappa_\alpha}\hat{\alpha}_i^{1-\kappa_\alpha}, \quad \kappa_\alpha > 1$$

$$\frac{L_i^l}{L_i} = 1 - (\frac{\hat{\alpha}_i}{\underline{\alpha}_i})^{-\kappa_\alpha}.$$

Observe that $\frac{W_i^h L_i^h}{W_i^l L_i^l} = \frac{W_i^h L_i^h}{W_i^h \hat{\alpha}_i L_i^l} = \frac{\frac{\kappa_\alpha}{\kappa_\alpha - 1}(\frac{\hat{\alpha}_i}{\underline{\alpha}_i})^{-\kappa_\alpha}}{1-(\frac{\hat{\alpha}_i}{\underline{\alpha}_i})^{-\kappa_\alpha}}$, which is an increasing function of $(\frac{\hat{\alpha}_i}{\underline{\alpha}_i})^{-\kappa_\alpha}$, the share of workers who become a researcher.

Derive the demand for researchers and aggregate knowhow. From equation (SA.B.4),under zero fixed offshore R&D cost and frictionless multinational production ($\phi_{oim}^P = 1$), we have

$$\begin{aligned} L_i^h &\propto (\frac{1}{W_i^h})^{\frac{1}{1-\gamma}} \sum_o E_o \cdot (\phi_{oi}^R)^{1-\gamma} \int_{\hat{z}_{oi}^R}^\infty (\tilde{z}^R)^{\frac{1}{1-\gamma}} [\int_0^\infty (z^P)^{\frac{\theta}{1-\gamma}} dG^P(z^P|\tilde{z}^R\phi_{oi}^R)] dG_o^E(\tilde{z}^R) \\ &\equiv (\frac{1}{W_i^h})^{\frac{1}{1-\gamma}} \sum_o Z_{oi} \\ &\equiv (\frac{1}{W_i^h})^{\frac{1}{1-\gamma}} Z_i \end{aligned} \tag{SA.B.6}$$

where $Z_{oi} \equiv E_o \cdot (\phi_{oi}^R)^{1-\gamma} \int_{\hat{z}_{oi}^R}^\infty (\tilde{z}^R)^{\frac{1}{1-\gamma}} [\int_0^\infty (z^P)^{\frac{\theta}{1-\gamma}} dG^P(z^P|\tilde{z}^R\phi_{oi}^R)] dG_o^E(\tilde{z}^R)$ is the total stock of knowhow in country i among firms from o, which is an increasing function of ϕ_{oi}^R, and $Z_i \equiv \sum_o Z_{oi}$.

Equation (SA.B.6) implies that the total R&D expenditures in country d are $\propto (Z_d)^{1-\gamma}(L_d^h)^\gamma$. Thus we have the following labor market clearing condition for $d = i, i'$:

$$
\begin{aligned}
W_d^l L_d^l &\propto \lambda_d \cdot \sum_{d'} X_{d'} = \frac{(T_d/W_d^l)^\theta}{\sum_{m'}(T_{m'}/W_{m'}^l)^\theta} \cdot \sum_{d'} X_{d'} \\
W_d^h L_d^h &\propto \frac{(Z_d)^{1-\gamma}(L_d^h)^\gamma}{\sum_{d'=1}^N (Z_{d'})^{1-\gamma}(L_{d'}^h)^\gamma} \cdot (\sum_{d'} X_{d'}),
\end{aligned}
$$

which implies

$$
\begin{aligned}
\frac{W_i^l L_i^l}{W_{i'}^l L_{i'}^l} &= \frac{(T_i/W_i^l)^\theta}{(T_{i'}/W_{i'}^l)^\theta} \\
\frac{W_i^h L_i^h}{W_{i'}^h L_{i'}^h} &= \frac{(Z_i)^{1-\gamma}(L_i^h)^\gamma}{(Z_{i'})^{1-\gamma}(L_{i'}^h)^\gamma}.
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
\frac{L_i^l}{L_{i'}^l} &= \frac{(\frac{T_i}{W_i^l})^\theta / W_i^l}{(\frac{T_{i'}}{W_{i'}^l})^\theta / W_{i'}^l} = \frac{(\frac{T_i}{W_i^h \cdot \hat{\alpha}_i})^\theta / W_i^l}{(\frac{T_{i'}}{W_{i'}^h \cdot \hat{\alpha}_{i'}})^\theta / W_{i'}^l} = \frac{(T_i)^\theta / (W_i^h \cdot \hat{\alpha}_i)^{1+\theta}}{(T_{i'})^\theta / (W_{i'}^h \cdot \hat{\alpha}_{i'})^{1+\theta}} \qquad \text{(SA.B.7)} \\
\Leftrightarrow \frac{1-(\frac{\hat{\alpha}_i}{\underline{\alpha}_i})^{-\kappa_\alpha}}{1-(\frac{\hat{\alpha}_{i'}}{\underline{\alpha}_{i'}})^{-\kappa_\alpha}} &= \frac{(T_i)^\theta}{(T_{i'})^\theta} \cdot \left(\frac{Z_{i'}}{Z_i}\right)^{(1+\theta)(1-\gamma)} \cdot \left(\frac{L_i}{L_{i'}}\right)^{(1+\theta)(1-\gamma)-1} \cdot \left(\frac{\underline{\alpha}_i}{\underline{\alpha}_{i'}}\right)^{\kappa_\alpha(1-\gamma)(1+\theta)} \cdot \left(\frac{\hat{\alpha}_{i'}}{\hat{\alpha}_i}\right)^{[1+(\kappa_\alpha-1)(1-\gamma)](1+\theta)} \\
\Leftrightarrow \frac{[1-(\frac{\hat{\alpha}_i}{\underline{\alpha}_i})^{-\kappa_\alpha}](\frac{\hat{\alpha}_i}{\underline{\alpha}_i})^{[1+(\kappa_\alpha-1)(1-\gamma)](1+\theta)}}{[1-(\frac{\hat{\alpha}_{i'}}{\underline{\alpha}_{i'}})^{-\kappa_\alpha}](\frac{\hat{\alpha}_{i'}}{\underline{\alpha}_{i'}})^{[1+(\kappa_\alpha-1)(1-\gamma)](1+\theta)}} &= \frac{(T_i)^\theta}{(T_{i'})^\theta} \cdot \left(\frac{Z_{i'}}{Z_i}\right)^{(1+\theta)(1-\gamma)} \cdot \left(\frac{L_i}{L_{i'}}\right)^{(1+\theta)(1-\gamma)-1} \cdot \left(\frac{\underline{\alpha}_i}{\underline{\alpha}_{i'}}\right)^{-\gamma(1+\theta)}
\end{aligned}
$$

Note that $[1-x^{-\kappa_\alpha}]x^{[1+(\kappa_\alpha-1)(1-\gamma)](1+\theta)}$ increases monotonically in x. Therefore the $RHS > 1$ if and only if $\frac{\hat{\alpha}_i}{\underline{\alpha}_i} > \frac{\hat{\alpha}_{i'}}{\underline{\alpha}_{i'}}$, i.e., i' has a higher share of workers in innovation, which also implies a higher share of labor income in i is from R&D than in i'.

□

SA.B.4 Proof of Proposition 3

Proof. In this proof, we use $\dot{x}$ to denote variables under frictionless MP with offshore R&D; $\tilde{x}$ to denote variables under frictionless MP without offshore R&D; $\hat{x}$ to denote variables with offshore R&D but no MP; and x (without accent) to denote variables with neither offshore R&D nor MP. These cases are illustrated in the matrix below:

	No Offshore R&D	With Offshore R&D
Frictionless MP	$\tilde{x}$	$\dot{x}$
No MP	x	$\hat{x}$

Also note that when MP is infeasible, each country has the same of labor in come from R&D versus from production, $\frac{W_i^h L_i^h}{L_i^l W_i^l} = \frac{\hat{W}_{i'}^h \hat{L}_{i'}^h}{\hat{L}_{i'}^l \hat{W}_{i'}^l} = \frac{\gamma}{\sigma-1}$.

Case 1: i' initially specializes in innovation. Suppose when MP is frictionless and offshore R&D is prohibitively costly, country i specializes in production and country i' specializes in innovation.

We first prove that $\frac{\dot{W}_{i'}^h \dot{L}_{i'}^h}{\dot{L}_{i'}^l \dot{W}_{i'}^l} > \frac{\hat{W}_{i'}^h \hat{L}_{i'}^h}{\hat{L}_{i'}^l \hat{W}_{i'}^l}$, i.e., the increase in $\phi_{ii'}^R$ leads to greater specialization of i' in innovation when MP is frictionless than when MP is infeasible.

Consider increasing $\phi^R_{ii'}$ when MP is frictionless. In this case, the specialization pattern is characterized by equation (SA.B.7). After an increase in $\phi^R_{ii'}$, the total access to knowhow in i' increases, so $\frac{\dot{Z}_{i'}}{\dot{Z}_i} > \frac{\tilde{Z}_{i'}}{\tilde{Z}_i}$. Following equation (SA.B.7), then, this implies that country i' becomes more specialized in innovation relative to i. As the specialization level of the entire economy is a constant, i' becoming more specialized relative to i means i' is more specialized in the 'dot' economy than in the 'tilde' economy, i.e., $\frac{\dot{W}^h_{i'}\dot{L}^h_{i'}}{\dot{L}^l_{i'}\dot{W}^l_{i'}} > \frac{\tilde{W}^h_{i'}\tilde{L}^h_{i'}}{\tilde{L}^l_{i'}\tilde{W}^l_{i'}}$. Since in the absence of offshore R&D, i' specializes in innovation, we have $\frac{\tilde{W}^h_{i'}\tilde{L}^h_{i'}}{\tilde{L}^l_{i'}\tilde{W}^l_{i'}} > \frac{\hat{W}^h_{i'}\hat{L}^h_{i'}}{\hat{L}^l_{i'}\hat{W}^l_{i'}} = \frac{\gamma}{\sigma-1}$, which means $\frac{\dot{W}^h_{i'}\dot{L}^h_{i'}}{\dot{L}^l_{i'}\dot{W}^l_{i'}} > \frac{\hat{W}^h_{i'}\hat{L}^h_{i'}}{\hat{L}^l_{i'}\hat{W}^l_{i'}}$. Under the Pareto assumption (see the discussion under equation (SA.B.5)), this implies a higher share of workers engaged in R&D in the 'dot' economy than in the 'hat' economy.

Recall that from equation (B.14), I_{oi} denote the payment to researchers in i for product development by firms from o. Since there is no fixed offshore R&D cost, we have $(\frac{\dot{Z}_{ii'}}{\sum_{o=i,i'}\dot{Z}_{oi}}) = (\frac{\hat{Z}_{ii'}}{\sum_{o=i,i'}\hat{Z}_{oi}})$. Noting that

$$\begin{aligned}\frac{\dot{I}_{ii'}}{\dot{I}_{ii'}+\dot{I}_{i'i'}} &= (\frac{\dot{Z}_{ii'}}{\sum_{o=i,i'}\dot{Z}_{oi}}) \\ \frac{\hat{I}_{ii'}}{\hat{I}_{ii'}+\hat{I}_{i'i'}} &= (\frac{\hat{Z}_{ii'}}{\sum_{o=i,i'}\hat{Z}_{oi}}),\end{aligned} \tag{SA.B.8}$$

It follows that more workers in i' work in R&D at firms from i in the 'dot' economy than in the 'hat' economy. Because in these two economies, the total knowhow available to researchers are the same, more varieties are developed at foreign affiliates in the 'dot' economy than in the 'hat' economy.

Case 2: i' initially specializes in production. Suppose when MP is frictionless and offshore R&D is prohibitively costly, country i specializes in innovation and country i' specializes in production.

This means $\frac{\tilde{W}^h_{i'}\tilde{L}^h_{i'}}{\tilde{L}^l_{i'}\tilde{W}^l_{i'}} < \frac{\gamma}{\sigma-1} = \frac{\hat{W}^h_{i'}\hat{L}^h_{i'}}{\hat{L}^l_{i'}\hat{W}^l_{i'}}$.

As in the first case, when MP is frictionless, specialization is given determined by equation (SA.B.7). An increase in $\phi^R_{ii'}$ means $\frac{\dot{Z}_{i'}}{\dot{Z}_i} > \frac{\tilde{Z}_{i'}}{\tilde{Z}_i}$, which implies that after the change, country i' become more specialized in innovation (and less specialized in production) relative to i than before. As the specialization level of the entire economy is a constant, i' becoming more specialized in innovation relative to i means i' is more specialized in innovation in the 'dot' economy than in the 'tilde' economy, i.e., $\frac{\dot{W}^h_{i'}\dot{L}^h_{i'}}{\dot{L}^l_{i'}\dot{W}^l_{i'}} > \frac{\tilde{W}^h_{i'}\tilde{L}^h_{i'}}{\tilde{L}^l_{i'}\tilde{W}^l_{i'}}$.

As the change in $\phi^R_{ii'}$ is small, from the continuity of the model, $\frac{\dot{W}^h_{i'}\dot{L}^h_{i'}}{\dot{W}^l_{i'}\dot{L}^l_{i'}}$ would still be smaller than $\frac{\gamma}{\sigma-1}$, and we have

$$\frac{\tilde{W}^h_{i'}\tilde{L}^h_{i'}}{\tilde{L}^l_{i'}\tilde{W}^l_{i'}} < \frac{\dot{W}^h_{i'}\dot{L}^h_{i'}}{\dot{L}^l_{i'}\dot{W}^l_{i'}} < \frac{\gamma}{\sigma-1} = \frac{\hat{W}^h_{i'}\hat{L}^h_{i'}}{\hat{L}^l_{i'}\hat{W}^l_{i'}} = \frac{W^h_{i'}L^h_{i'}}{L^l_{i'}W^l_{i'}}.$$

Under the Pareto assumption (see the discussion under equation (SA.B.5)), this implies a lower share of workers engaged in R&D in the 'dot' economy than in the 'hat' economy.

Noting that equation (SA.B.8) continues to hold in this case, so among researchers in i', the same fraction work at foreign affiliates in the 'dot' and 'hat' economy. A smaller number of worker in research in the 'dot' economy means a smaller number of researchers work at foreign affiliates. Since the available foreign knowhow is the same in 'dot' and 'hat' economies, it follows that fewer varieties are developed at foreign affiliates in the 'dot' economy than in the 'hat' economy.

□

SA.B.5 Proof of Proposition 4

Proof. I proceed in four steps. The first two steps express the real wage for low-skill workers as flow variables. The third step derives the relationship between real wage and real income. The fourth and final step combines results from the first three steps to derive $\frac{X_d}{P_d}$ as a function of several ratios and calculate the gains from openness by setting some of the ratios to their autarky values, which is usually 1 or a model constant. I then compare my expression of the gains from openness to that in the literature.

Step 1: Expressing real wage for low-skill workers as flow variables

The first step derives $\frac{W_d^l}{P_d}$, and is a bit tedious. To give a broad direction, I will express P_d as a function of the share of consumption expenditures of country d spent on goods *invented* in country d. Intuitively, P_d measures competitiveness of the product market in country d; for any given level of domestic invention, if a higher share of income is spent on these inventions, then it must be because country d has a poor access to varieties invented elsewhere, and P_d will thus be high. In a similar vein, I will express W_d^l as a function of the share of consumption expenditures of country d spent on the goods *produced* in country d. All else equal, if this share is higher, then the wage of country d is likely lower.

It will prove convenient to define K_{oi} as follows:

$$K_{oi} \equiv \int_0^\infty (z^P)^\theta V_{oi}(z^P) dz^P. \tag{SA.B.9}$$

Loosely speaking, K_{oi} is the productivity-adjusted number of varieties invented in i by firms from o.

From equation (B.15) we have the following:

$$P_d^{-\theta} = \theta(\frac{\sigma}{\sigma-1})^{-\theta} \frac{1}{\theta-(\sigma-1)} \Big(\frac{\sigma W_d^h f_d^M}{X_d}\Big)^{\frac{\theta-(\sigma-1)}{1-\sigma}} \sum_o \sum_i \tilde{\zeta}_{oid}^\theta \cdot K_{oi}$$

Similarly, from equation (B.14) we have:

$$X_{oid} = \theta(\frac{\sigma}{\sigma-1})^{-\theta} \frac{1}{\theta-(\sigma-1)} (\sigma W_d^h f_d^M)^{\frac{\theta-(\sigma-1)}{1-\sigma}} \Big(\frac{X_d}{P_d^{1-\sigma}}\Big)^{\frac{\theta}{\sigma-1}} \tilde{\zeta}_{oid}^\theta K_{oi}$$

Define $\lambda_{oid}^E = \frac{\sum_m X_{oimd}}{X_d}$, which denotes the share of expenditure in country d spent on goods invented in i by firms from country o, I obtain:

$$\lambda_{oid}^E = \frac{X_{oid}}{X_d} = \frac{\tilde{\zeta}_{oid}^\theta K_{oi}}{\sum_{o,i} \tilde{\zeta}_{oid}^\theta K_{oi}} = \frac{\tilde{\zeta}_{oid}^\theta K_{oi}}{P_d^{-\theta}} \cdot \theta(\frac{\sigma}{\sigma-1})^{-\theta} \frac{1}{\theta-(\sigma-1)} \Big(\frac{\sigma W_d^h f_d^M}{X_d}\Big)^{\frac{\theta-(\sigma-1)}{1-\sigma}},$$

which implies

$$P_d^{-\theta} = \frac{\theta(\frac{\sigma}{\sigma-1})^{-\theta} \frac{1}{\theta-(\sigma-1)} \Big(\frac{\sigma W_d^h f_d^M}{X_d}\Big)^{\frac{\theta-(\sigma-1)}{1-\sigma}} \cdot \tilde{\zeta}_{oid}^\theta K_{oi}}{\lambda_{oid}^E}.$$

Specializing this equation to $o = i = d$ gives

$$P_d^{-\theta} = \frac{\theta(\frac{\sigma}{\sigma-1})^{-\theta} \frac{1}{\theta-(\sigma-1)} \Big(\frac{\sigma W_d^h f_d^M}{X_d}\Big)^{\frac{\theta-(\sigma-1)}{1-\sigma}} \cdot \tilde{\zeta}_{ddd}^\theta K_{dd}}{\lambda_{ddd}^E} \tag{SA.B.10}$$

Now consider W_d^l. Define $\lambda_{dd}^T = \frac{\sum_{o,i} X_{oidd}}{X_d}$ as the fraction of country d's spending on goods produced in d, regardless where it is invented and which the headquarter countries of the firms are.

Noticing $\lambda^T_{dd} = \sum_{o,i} \psi_{oidd}\lambda^E_{oid}$ and recalling $\psi_{oimd} = \frac{\frac{1}{N}(\frac{T_m \phi^P_{oim}}{W^l_m \tau_{md}})^\theta}{\tilde{\zeta}^\theta_{oid}}$, we have

$$\lambda^T_{dd} = \sum_{o,i} \frac{\frac{1}{N}(\frac{T_d \phi^P_{oid}}{W^l_d \tau_{dd}})^\theta}{\tilde{\zeta}^\theta_{oid}} \lambda^E_{oid} \tag{SA.B.11}$$

$$\implies (W^l_d)^\theta = \frac{1}{\lambda^T_{dd}} \sum_{o,i} \frac{\frac{1}{N}(T_d \phi^P_{oid})^\theta}{\tilde{\zeta}^\theta_{oid}} \lambda^E_{oid}$$

Combining equations (SA.B.10) and (SA.B.11) gives

$$P_d^{-\theta} \cdot (W^l_d)^\theta = \frac{\theta(\frac{\sigma}{\sigma-1})^{-\theta} \frac{1}{\theta-(\sigma-1)} \left(\frac{\sigma W^h_d f^M_d}{X_d}\right)^{\frac{\theta-(\sigma-1)}{1-\sigma}}}{\lambda^T_{dd}} [\sum_{o,i} \frac{1}{N}(T_d \phi^P_{oid})^\theta \frac{\tilde{\zeta}^\theta_{ddd}}{\tilde{\zeta}^\theta_{oid}} \lambda^E_{oid}] \frac{K_{dd}}{\lambda^E_{ddd}} \tag{SA.B.12}$$

The term $[\sum_{o,i} \frac{1}{N}(T_d \phi^P_{oid})^\theta \frac{\tilde{\zeta}^\theta_{ddd}}{\tilde{\zeta}^\theta_{oid}} \lambda^E_{oid}]$ broadly captures the importance of country d as a production location, which can be derived as a function of flows as follows:

$$\frac{X_{oidd}}{X_d} = \lambda^E_{oid} \cdot \frac{\frac{1}{N}(\frac{T_d \phi^P_{oid}}{W^l_d \tau_{dd}})^\theta}{\tilde{\zeta}^\theta_{oid}} \implies T^\theta_d \frac{\sum_m X_{ddmd}}{X_{dddd}} \cdot \lambda^T_{dd} = \sum_{o,i} \lambda^E_{oid} \cdot \frac{\tilde{\zeta}^\theta_{ddd}}{\tilde{\zeta}^\theta_{oid}} \cdot \frac{1}{N}(T_d \phi^P_{oid})^\theta$$

Plugging this into equation (SA.B.12), I obtain:

$$\begin{aligned}
P_d^{-\theta} \cdot (W^l_d)^\theta &= \frac{\theta(\frac{\sigma}{\sigma-1})^{-\theta} \frac{1}{\theta-(\sigma-1)} \left(\frac{\sigma W^h_d f^M_d}{X_d}\right)^{\frac{\theta-(\sigma-1)}{1-\sigma}}}{\lambda^T_{dd}} \cdot T^\theta_d \cdot \frac{\sum_m X_{ddmd}}{X_{dddd}} \cdot \lambda^T_{dd} \cdot \frac{K_{dd}}{\lambda^E_{ddd}} \\
&\quad (\text{noting that } \lambda^E_{ddd} \equiv \frac{\sum_m X_{oimd}}{X_d} = \frac{\sum_{o,m} X_{odmd}}{X_d} \cdot \frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}}) \\
&= T^\theta_d \theta(\frac{\sigma}{\sigma-1})^{-\theta} \frac{1}{\theta-(\sigma-1)} \left(\frac{\sigma W^h_d f^M_d}{X_d}\right)^{\frac{\theta-(\sigma-1)}{1-\sigma}} \cdot \frac{\sum_m X_{ddmd}}{X_{dddd}} \cdot \frac{X_d}{\sum_{o,m} X_{odmd}} \cdot \frac{\sum_{o,m} X_{odmd}}{\sum_m X_{ddmd}} \cdot K_{dd} \\
&\quad (\text{noting that } T^\theta_d, \theta, \sigma, f^M_d \text{ are constants }) \\
\implies (\frac{W^l_d}{P_d}) &\propto (\frac{W^h_d}{X_d})^{\frac{\theta-\sigma+1}{(1-\sigma)\theta}} \cdot (\frac{X_{dddd}}{\sum_m X_{ddmd}})^{-\frac{1}{\theta}} \cdot (\frac{\sum_{o,m} X_{odmd}}{X_d})^{-\frac{1}{\theta}} \cdot (\frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}})^{-\frac{1}{\theta}} \cdot (K_{dd})^{\frac{1}{\theta}}.
\end{aligned}$$

We can already see that the real wage for low-skill workers is a function of $\frac{X_{dddd}}{\sum_m X_{ddmd}}, \frac{\sum_{o,m} X_{odmd}}{X_d}, \frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}}$. They capture the importance of foreign locations for production, the importance of varieties developed outside the country, and the importance of foreign firms in domestic R&D, respectively. Note also that all these ratios equal to one in autarky. Define $\hat{x}$ be the ratio between the baseline variable x and its value in autarky x', we have:

$$\frac{\widehat{W^l_d}}{P_d} = (\frac{\widehat{W^h_d}}{X_d})^{\frac{\theta-\sigma+1}{(1-\sigma)\theta}} (\frac{X_{dddd}}{\sum_m X_{ddmd}})^{-\frac{1}{\theta}} \cdot (\frac{\sum_{o,m} X_{odmd}}{X_d})^{-\frac{1}{\theta}} \cdot (\frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}})^{-\frac{1}{\theta}} \cdot (\widehat{K_{dd}})^{\frac{1}{\theta}} \tag{SA.B.13}$$

Step 2: deriving $\widehat{K_{dd}}$.

Now I derive $\widehat{K_{dd}} \equiv \frac{K_{dd}}{K'_{dd}}$, where K'_{dd} is the autarky value of K_{dd}.

Assume that the total number of high-skill workers in country i working directly on variety development is L^R_i and let L^R_{oi} be those working at R&D centers from o: $L^R_i = \sum_o L^R_{oi}$. Note that L^R_i is endogenous variable even though the total number of high-skill workers (which we denote by L_i^h) is assumed to be exogenous for this proposition.

$$
\begin{aligned}
L_{oi}^R &= R_{oi} \int_0^\infty [\int_0^\infty h_{oi}(z^P, z^R) g_{oi}^R(z^R) dz^R] \cdot g_{oi}^p(z^P) dz^P \\
&= \left(\frac{\gamma}{W_i^h}\right)^{\frac{1}{1-\gamma}} \cdot R_{oi} \cdot (\overline{\pi}_{oi}^P)^{\frac{1}{1-\gamma}} [\int_0^\infty (z^R)^{\frac{1}{1-\gamma}} g_{oi}^R(z^R) dz^R] \cdot [\int_0^\infty (z^P)^{\frac{\theta}{1-\gamma}} g_{oi}^P(z^P) dz^P],
\end{aligned}
$$

where I define $\overline{\pi}_{oi}^P$ as the component in $\overline{\pi}_{oi}(z^P)$ that is independent of z^P: $\overline{\pi}_{oi}^P \equiv \frac{\overline{\pi}_{oi}(z^P)}{(z^P)^\theta}$.

Summing across all origin countries:

$$
\begin{aligned}
L_i^R &= \left(\frac{\gamma}{W_i^h}\right)^{\frac{1}{1-\gamma}} \sum_o R_{oi} \cdot (\overline{\pi}_{oi}^P)^{\frac{1}{1-\gamma}} [\int_0^\infty (z^R)^{\frac{1}{1-\gamma}} g_{oi}^R(z^R) dz^R] \cdot [\int_0^\infty (z^P)^{\frac{\theta}{1-\gamma}} g_{oi}^P(z^P) dz^P] \qquad \text{(SA.B.14)} \\
\implies \left(\frac{\gamma}{W_i^h}\right)^{\frac{1}{1-\gamma}} &= \frac{L_i^R}{\sum_o R_{oi} \cdot (\overline{\pi}_{oi}^P)^{\frac{1}{1-\gamma}} [\int_0^\infty (z^R)^{\frac{1}{1-\gamma}} g_{oi}^R(z^R) dz^R] \cdot [\int_0^\infty (z^P)^{\frac{\theta}{1-\gamma}} g_{oi}^P(z^P) dz^P]}
\end{aligned}
$$

With this, now I characterize $\widehat{K}_{oi}$. By equation (SA.B.9):

$$
\begin{aligned}
K_{oi} &= \int_0^\infty (z^P)^\theta V_{oi}(z^P) dz^P \\
&= R_{oi} \cdot (\overline{\pi}_{oi}^P)^{\frac{\gamma}{1-\gamma}} \cdot \left(\frac{\gamma}{W_i^h}\right)^{\frac{\gamma}{1-\gamma}} [\int_0^\infty z^{R\frac{1}{1-\gamma}} g_{oi}^R(z^R) dz^R] \cdot [\int_0^\infty (z^P)^\theta \cdot (z^P)^{\frac{\theta\gamma}{1-\gamma}} g_{oi}^P(z^P) dz^P] \\
&\quad \text{(plugging in equation (SA.B.14))} \\
&= \frac{(L_i^R)^\gamma \cdot R_{oi} \cdot (\overline{\pi}_{oi}^P)^{\frac{\gamma}{1-\gamma}} \cdot [\int_0^\infty z^{R\frac{1}{1-\gamma}} g_{oi}^R(z^R) dz^R] \cdot [\int_0^\infty \cdot (z^P)^{\frac{\theta}{1-\gamma}} g_{oi}^P(z^P) dz^P]}{(\sum_o R_{oi} \cdot (\pi_{oi}^P)^{\frac{1}{1-\gamma}} [\int_0^\infty (z^R)^{\frac{1}{1-\gamma}} g_{oi}^R(z^R) dz^R] \cdot [\int_0^\infty (z^P)^{\frac{\theta}{1-\gamma}} g_{oi}^P(z^P) dz^P])^\gamma} \\
&= (L_i^R)^\gamma \cdot \left(R_{oi} \cdot [\int_0^\infty z^{R\frac{1}{1-\gamma}} g_{oi}^R(z^R) dz^R] \cdot [\int_0^\infty \cdot (z^P)^{\frac{\theta}{1-\gamma}} g_{oi}^P(z^P) dz^P]\right)^{1-\gamma} \cdot (\frac{I_{oi}}{\sum_o I_{oi}})^\gamma,
\end{aligned}
$$

where the last equation follows from that all researchers in a country are paid the same wage and the definition of I_{oi}.

Specializing the above equation to $o = i = d$, noting that because domestic firms always do R&D locally, $\left(R_{dd} \cdot [\int_0^\infty z^{R\frac{1}{1-\gamma}} g_{dd}^R(z^R) dz^R] \cdot [\int_0^\infty \cdot (z^P)^{\frac{\theta}{1-\gamma}} g_{dd}^P(z^P) dz^P]\right)^{1-\gamma}$ is a constant that does not respond to economic shocks and that in autarky, $\frac{I_{dd}}{\sum_o I_{od}} = 1$, we have

$$
\begin{aligned}
\widehat{K}_{dd} &= (\widehat{L_d^R})^\gamma \cdot (\frac{I_{dd}}{\sum_o I_{od}})^\gamma \qquad \text{(SA.B.15)} \\
&= (\frac{L_d^R}{L_d^{R'}})^\gamma \cdot (\frac{I_{dd}}{\sum_o I_{od}})^\gamma \\
&= (\frac{L_d^R / L_d^h}{L_d^{R'} / L_d^{h'}})^\gamma \cdot (\frac{I_{dd}}{\sum_o I_{od}})^\gamma,
\end{aligned}
$$

in which $L_d^{R'}$ is the number of high-skill workers in country d working on variety development in autarky, and $L_d^{h'}$ is total high-skill workers in d in autarky. Since in autarky a fixed share of income is given to marketing and moreover, there is no fixed overhead for offshore R&D (as domestic firms do not pay the fixed R&D center setup cost), it follows from equation (B.14) that $L_d^{R'} / L_d^{h'} = \frac{\gamma(\sigma-1)}{\theta-(1-\gamma)(\sigma-1)}$. The only item in $\widehat{K}_{dd}$ that is yet to be characterized is thus L^R_d / L^h_d. To this end, note that

$$\frac{L^R_d}{L^h_d} = \frac{\sum_o I_{od}}{\sum_o I_{od} + \sum_o F^M_{od} + \sum_{o\neq d} F^R_{od}}$$
$$(\text{using } F^M_{od} = \frac{1}{\sigma}\cdot\Big(\frac{\theta-(\sigma-1)}{\theta}\Big)\sum_{i,m} X_{oimd})$$
$$= \frac{\sum_o I_{od}}{\sum_o I_{od} + \frac{\theta-(\sigma-1)}{\theta\sigma} X_d + \sum_{o\neq d} F^R_{od}}$$

Note also that for $o \neq d$, given the assumption of $G^R_o(\tilde{z}^R)$ being a Pareto distribution, I can write the total fixed cost paid by firms from o opening up R&D centers in d as a function of total profit generated by these R&D centers. Letting $\hat{z}^R_{od}$ be the cutoff for firms to conduct offshore R&D in d:

$$F^R_{od} = f^R_{od} W^h_d E_o \int_{\hat{z}^R_{od}}^{\inf} dG^R_o(\tilde{z}^R) = f^R_{od} W^h_d E_o (\frac{\hat{z}^R_{oi}}{\underline{Z}^R_o})^{-\kappa_R}.$$

The indifference condition at the cutoff implies:

$$\pi_{oc}(\hat{z}^R_{od}) = \int \pi_{od}(\hat{z}^R_{od}\phi^R_{od}, z^P) dG^P_{od}(z^P)$$
$$= (\hat{z}^R_{od}\phi^R_{od})^{\frac{1}{1-\gamma}} \cdot \underbrace{(\gamma^{\frac{\gamma}{1-\gamma}} - \gamma^{\frac{1}{1-\gamma}})\Big(\frac{1}{W^h_d}\Big)^{\frac{\gamma}{1-\gamma}} [\sum_{d'} \frac{(\sigma-1)^{1+\theta}}{\theta-(\sigma-1)} \sigma^{\frac{\sigma\theta}{1-\sigma}} (f^M_{d'} W^h_{d'})^{\frac{\sigma-1-\theta}{\sigma-1}} X_{d'}^{\frac{\theta}{\sigma-1}} P^\theta_{d'} (\tilde{\zeta}_{oid'})^\theta]^{\frac{1}{1-\gamma}} \int_0^\infty (z^P)^{\frac{\theta}{1-\gamma}} dG^P_{od}(z^P)}_{\equiv Z_{od}}$$
$$\equiv (\hat{z}^R_{od}\phi^R_{od})^{\frac{1}{1-\gamma}} \cdot Z_{od}$$
$$= f^R_{od} W^h_d,$$

in which Z_{oi} is introduced to shorten the expressions. This gives:

$$(\hat{z}^R_{od})^{\frac{1}{1-\gamma}} (\phi^R_{od})^{\frac{1}{1-\gamma}} = \frac{f^R_{od} W^h_d}{Z_{od}}$$

Total variable profit from od is

$$\begin{aligned}
\Pi_{od} &= E_o \int_{\hat{z}_{od}\phi^R_{od}}^\infty \pi_{od}(z^R) dG^R_o(\tilde{z}^R) \qquad \text{(SA.B.16)}\\
&= E_o \cdot Z_{od} \cdot \int_{\hat{z}^R_{od}}^\infty (z^R_{oi}\phi^R_{od})^{\frac{1}{1-\gamma}} dG^R_o(\tilde{z}^R)\\
&= E_o \cdot f^R_{od} W^h_d \cdot (\frac{\hat{z}^R_{od}}{\underline{Z}^R_o})^{-\kappa^R} \cdot \frac{\kappa^R}{\kappa^R - \frac{1}{1-\gamma}}\\
&= F^R_{od} \cdot \frac{\kappa^R}{\kappa^R - \frac{1}{1-\gamma}}
\end{aligned}$$

i.e., we obtain the fraction of fixed RD cost as a share of profit as: $\frac{\kappa^R - \frac{1}{1-\gamma}}{\kappa^R}$, which is assumed to be below 1 for profit to be integrable.

Plugging this into equation (SA.B.15), I obtain

$$\widehat{K}_{dd} = (\frac{I_{dd}}{\sum_o I_{od}})^{\gamma} \cdot \Big(\frac{\sum_o I_{od}}{\sum_o I_{od} + \frac{\kappa^R - \frac{1}{1-\gamma}}{\kappa^R} \sum_{o\neq d} \Pi_{od} + \frac{\theta-(\sigma-1)}{\theta\sigma} X_d} \cdot \frac{\theta - (1-\gamma)(\sigma-1)}{\gamma(\sigma-1)}\Big)^{\gamma} \tag{SA.B.17}$$
$$= (\frac{I_{dd}}{\sum_o I_{od}})^{\gamma} \cdot \Big(\frac{1}{1 + \frac{\kappa^R - \frac{1}{1-\gamma}}{\kappa^R} \cdot \frac{1-\gamma}{\gamma} \cdot \frac{\sum_{o\neq d} I_{od}}{\sum_o I_{od}} + \frac{\theta-(\sigma-1)}{\theta\sigma} \frac{X_d}{\sum_o I_{od}}} \cdot \frac{\theta - (1-\gamma)(\sigma-1)}{\gamma(\sigma-1)}\Big)^{\gamma},$$

in other words, $\widehat{K}_{dd}$ is only a function of 1) the fraction of R&D done by firms from d, $\frac{I_{dd}}{\sum_o I_{od}}$, and 2), the fraction of variable R&D expenses in income: $\frac{\sum_o I_{od}}{X_d}$.

Step 3: deriving $\widehat{\frac{W_d^h}{X_d}}$ and $\widehat{\frac{W_d^l}{X_d}}$. I now deriving ratios between baseline and autarky values for $\frac{W_d^h}{X_d}$ and $\frac{W_d^l}{X_d}$. Note that because the supply of high- and low-skill workers are exogenous, $\widehat{\frac{W_d^h}{X_d}} = \widehat{\frac{W_d^h L_d^h}{X_d}}$ and $\widehat{\frac{W_d^l}{X_d}} = \widehat{\frac{W_d^l L_d^l}{X_d}}$, i.e., we only need to derive the change in income share of high- and low-skill workers.

From equations (B.17) and (SA.B.16), we have $\frac{W_d^h L_d^h}{X_d} = \frac{\sum_o I_{od} + \frac{\kappa^R - \frac{1}{1-\gamma}}{\kappa^R} \cdot \frac{1-\gamma}{\gamma} \cdot \sum_{o\neq d} I_{od} + \frac{\theta-(\sigma-1)}{\theta\sigma} X_d}{X_d}$, In autarky, this ratio collapses to $\frac{\theta-(1-\gamma)(\sigma-1)}{\sigma\theta}$, so we have

$$\widehat{\frac{W_d^h}{X_d}} = \frac{\sigma\theta}{\theta-(1-\gamma)(\sigma-1)} \cdot [\frac{\sum_o I_{od}}{X_d} + \frac{\kappa^R - \frac{1}{1-\gamma}}{\kappa^R} \cdot \frac{1-\gamma}{\gamma} \cdot \frac{\sum_{o\neq d} I_{od}}{X_d} + \frac{\theta-(\sigma-1)}{\theta\sigma}]. \tag{SA.B.18}$$

From equation (B.17), we have $\frac{W_d^l L_d^l}{X_d} = \frac{Y_d}{X_d}$. In autarky, this ratio is simply $\frac{\sigma-1}{\sigma}$, so

$$\widehat{\frac{W_d^l}{X_d}} = \frac{\sigma}{\sigma-1}\frac{Y_d}{X_d}. \tag{SA.B.19}$$

Step 4: Putting all together

Combining equations (SA.B.13), (SA.B.17), (SA.B.18), and (SA.B.19), I obtain:

$$\widehat{\frac{X_d}{P_d}} = (\widehat{\frac{W_d^l}{P_d}}) \times (\widehat{\frac{X_d}{W_d^l}}) \tag{SA.B.20}$$
$$= (\widehat{\frac{W_d^h}{X_d}})^{\frac{\theta-\sigma+1}{(1-\sigma)\theta}} \times (\frac{X_{dddd}}{\sum_m X_{ddmd}})^{-\frac{1}{\theta}} \times (\frac{\sum_{o,m} X_{odmd}}{X_d})^{-\frac{1}{\theta}} \times (\frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}})^{-\frac{1}{\theta}} \times (\widehat{K_{dd}})^{\frac{1}{\theta}} \times (\widehat{\frac{X_d}{W_d^l}})$$
$$= (\frac{X_{dddd}}{\sum_m X_{ddmd}})^{-\frac{1}{\theta}} \times (\frac{\sum_{o,m} X_{odmd}}{X_d})^{-\frac{1}{\theta}} \times (\frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}})^{-\frac{1}{\theta}} \times (\frac{I_{dd}}{\sum_o I_{od}})^{\frac{\gamma}{\theta}}$$
$$\times \Big(\frac{1}{1 + \frac{\kappa^R - \frac{1}{1-\gamma}}{\kappa^R} \cdot \frac{1-\gamma}{\gamma} \cdot \frac{\sum_{o\neq d} I_{od}}{\sum_o I_{od}} + \frac{\theta-(\sigma-1)}{\theta\sigma} \frac{X_d}{\sum_o I_{od}}} \cdot \frac{\theta - (1-\gamma)(\sigma-1)}{\gamma(\sigma-1)}\Big)^{\frac{\gamma}{\theta}}$$
$$\times \Big(\frac{\sigma\theta}{\theta-(1-\gamma)(\sigma-1)} \cdot [\frac{\sum_o I_{od}}{X_d} + \frac{\kappa^R - \frac{1}{1-\gamma}}{\kappa^R} \cdot \frac{1-\gamma}{\gamma} \cdot \frac{\sum_{o\neq d} I_{od}}{X_d} + \frac{\theta-(\sigma-1)}{\theta\sigma}]\Big)^{\frac{\theta-\sigma+1}{(1-\sigma)\theta}} \times \frac{X_d}{Y_d} \times \frac{\sigma-1}{\sigma}$$

$$\equiv (\frac{X_{dddd}}{\sum_m X_{ddmd}})^{-\frac{1}{\theta}} \times (\frac{\sum_{o,m} X_{odmd}}{X_d})^{-\frac{1}{\theta}} \times (\frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}})^{-\frac{1}{\theta}} \times (\frac{I_{dd}}{\sum_o I_{od}})^{\frac{\gamma}{\theta}} \times \frac{X_d}{Y_d} \times f(\frac{\sum_o I_{od}}{X_d}, \frac{I_{dd}}{\sum_o I_{od}}),$$

in which $f(\frac{\sum_o I_{od}}{X_d}, \frac{I_{dd}}{\sum_o I_{od}})$ is defined to be a function of only model elasticities and two measures, $\frac{\sum_o I_{od}}{X_d}$ and

$\frac{I_{dd}}{\sum_o I_{od}}$.[4]

This completes the proof of Proposition 1. This proof holds when z^R and z^P are independent. But accommodate any arbitrary distributions for $G^P_{oi}(z^P)$ and flexible Pareto distributions for $G^R_o(\tilde{z}^R)$. $\square$

SA.C Quantification

SA.C.1 Numerical Implementation

Table 5 of the text summarizes the three categories of parameters and the corresponding moments that identify them. I design a nested fixed point algorithm based on the feature of this problem to pin down the parameters efficiently. Before explaining the algorithm, it is useful to review the conditions characterizing the competitive equilibrium.

As discussed in Online Appendix Section B, equations (B.15), (B.16), (B.17), and (B.18) jointly characterize a fixed point system in wages and occupation choice $\{W^h_d, W^l_d, \hat{a}_d : d = 1, ...N\}$, prices $\{P_d : d = 1, ...N\}$, and aggregate expenditures $\{X_d : d = 1, ...N\}$. Furthermore, the moments that identify (subject to normalization) manufacturing TFP, $\{T_m | m = 1, ..., N\}$, measure of firms, $\{E_o | o = 1, .., N\}$, and host fixed effects in offshore activities $\{\phi^P_m, \phi^R_i, \phi^{fR}_i | i, m = 1, .., N\}$ are characterized by the following system of equations:

$$\begin{aligned}
T_m &: \frac{X_m}{P_m} = \widehat{\text{Real GDP}}_m, \; m = 1, ...N \qquad \text{(SA.C.1)}\\
E_o &: \frac{\sum_i I_{oi}}{\sum_{o,i} I_{oi}} = \frac{\widehat{\sum_i I_{oi}}}{\sum_{o,i} I_{oi}}, \; o = 1, ...N\\
\phi^P_m &: \frac{\sum_{o \neq m} Y_{om}}{\sum_o Y_{om}} = \frac{\widehat{\sum_{o \neq m} Y_{om}}}{\sum_o Y_{om}}, \; m = 1, ...N\\
\phi^R_i &: \frac{\sum_{o \neq i} I_{oi}}{\sum_o I_{oi}} = \frac{\widehat{\sum_{o \neq i} I_{oi}}}{\sum_o I_{oi}}, \; i = 1, ...N\\
\phi^{fR}_i &: \frac{\sum_{o \neq i} R_{oi}}{\sum_o R_{oi}} = \frac{\widehat{\sum_{o \neq i} R_{oi}}}{\sum_o R_{oi}}, \; i = 1, ...N
\end{aligned}$$

The the right-hand sides of these equations are the data. The left-hand side are their model counterparts. I write in front of each equation a fundamental variable (e.g., T_m) to stress that the model predictions are a function of these fundamentals.

Together with equations (B.15), (B.16), (B.17), and (B.18), equation (SA.C.1) characterizes a fixed point in the model fundamentals and endogenous outcomes, such that: 1) the solution to the fixed point problem is a competitive equilibrium; 2), the solution to the fixed point problem ensures that the model matches the data exactly as specified in ($SA.C.1$). I implement the following algorithm.

1. Choose $\underline{z}^P_H$ and κ^P.

   (a) Choose the 17 parameters governing geographic frictions: $\{\overrightarrow{\beta^{P,om}}, \overrightarrow{\beta^{P,im}}, \overrightarrow{\beta^R}, \overrightarrow{\beta^{cR}}\}$ and s

      i. Solve equations (B.15), (B.16), (B.17), (B.18), and (SA.C.1) jointly for the following: $\{W^h_d, W^l_d, \hat{a}_d : d = 1, ...N\}$, $\{P_d : d = 1, ...N\}$,$\{X_d : d = 1, ...N\}$, $\{T_m | m = 1, ..., N\}$, $\{E_o | o = 1, .., N\}$, and $\{\phi^P_m, \phi^R_i, \phi^{fR}_i | i, m = 1, .., N\}$.

[4] $f(\frac{\sum_o I_{od}}{X_d}, \frac{I_{dd}}{\sum_o I_{od}})$ collects the remaining terms and can be rearranged to be:

$$f(\tfrac{\sum_o I_{od}}{X_d}, \tfrac{I_{dd}}{\sum_o I_{od}}) = [1 + \tfrac{(1-\gamma)\kappa^R - 1}{\gamma \kappa^R}(1 - \tfrac{I_{dd}}{\sum_o I_{od}}) + \tfrac{\theta - (\sigma-1)}{\theta\sigma} \tfrac{X_d}{\sum_o I_{od}}]^{\frac{1-\gamma}{\theta} - \frac{1}{\sigma-1}} \left(\tfrac{\sum_o I_{od}}{X_d}\right)^{\frac{1}{\theta} - \frac{1}{\sigma-1}} \underbrace{[\theta - (1-\gamma)(\sigma-1)]^{\frac{1}{\sigma-1} - \frac{1-\gamma}{\theta}} \gamma^{-\frac{\gamma}{\theta}} (\sigma-1)^{1-\frac{\gamma}{\theta}} \theta^{\frac{1}{\theta} - \frac{1}{\sigma-1}} \sigma^{\frac{1}{\theta} - \frac{1}{\sigma-1} - 1}}_{\text{constant}}$$

ii. Simulate 5e4 firms. I assign the number of firms from a country to be proportional to its size and draw $\tilde{z}^R$ for these firms from its calibrated knowhow distribution.

iii. Solve for the optimal offshore R&D and production decision of these firms. Then estimate the same specifications as in the left panel of Table C.4 using the model-simulated data.

iv. Evaluate the objective function $f = \sum_{k=1}^{22}(\frac{x_k - \widehat{x}_k}{\widehat{\sigma}_k})^2$, where $x_k, k = 1, ..., 22$ is a model-based regression coefficient, $\widehat{x}_k$ is the empirical estimate, and $\hat{\sigma}_k$ is the standard error of $\hat{x}_k$.

(b) If the choice of $\{\overrightarrow{\beta^{P,om}}, \overrightarrow{\beta^{P,im}}, \overrightarrow{\beta^{R}}, \overrightarrow{\beta^{cR}}\}$ and s minimize f defined above, proceed to step 2, otherwise return to Step 1.(a) and try a different set of parameter values.

2. Compare the model-based firm size distribution to its data counterparts (Panel A of Table 5). If they are close enough, exit; otherwise return to Step 1.

Additional details on implementing the above algorithm. First, in searching over the space of the 17 geographic parameters, I try multiple starting points using two algorithms implemented by Knitro, interior point and active-set. Both algorithms give similar results.

Second, in solving for the fixed point problem in Step 1.(a).i, for a given set of fundamentals and aggregate prices and wages, I evaluate the endogenous objects in equations (B.15), (B.16), (B.17), (B.18), and (SA.C.1). These objects can be found by sequentially calculating equations (B.11), (B.12), (B.13), and (B.14). Most of these expressions are analytical and hence can be directly evaluated. For the ones that cannot be analytically evaluated, I approximate their values numerically as follows. First, the cutoff for offshore R&D, $\hat{z}^R_{oi}$ is given by the implicit function in equation (B.12). I analytically integrate over $\pi_{oi}(z^P, z^R)$ to obtain function $\pi^R_{oi}(z^R)$,[5] and then determine the cutoff $\hat{z}^R_{oi}$ as the indifference point for offshore R&D $\pi^R_{oi}(\tilde{z}^R \phi^R_{oi}) = f^R_{oi} W^h_i$ using the Brent method. Second, $V_{oi}(z^P)$ in equation (B.14) is a an integration of $v_{oi}(z^P, z^R)$ over the $\mathbb{Z}^R$ space; moreover, $V_{oi}(z^P)$ itself becomes an integrand for X_{oid} and P_d. For this step analytical integration is not available. I approximate for $V_{oi}(z^P)$ numerically using an adaptive Cash-Karp algorithm. The number of such numerical approximations for each evaluation of equations (B.15), (B.17), (B.18), (B.16), and (SA.C.1) increases quadratically with the number of countries in the sample; solving the equation systems and then finding the best-fit geographic parameters requires thousands of such evaluations. Step 1.(a).i is implemented in C++ to speed up the computation.

SA.C.2 Decomposition of R&D and Sources of Firm Profit in the Calibrated Economy

The model provides a measurement device for the role of offshore R&D in firms' global organization of production, and the contribution of offshore R&D to national income. Table SA.C.1 summarizes the measurement according to the baseline calibration.

Distribution of manufacturing by R&D modes. Columns 1 through 4 of Table SA.C.1 decompose R&D by its source (whether done by local firms) and use (whether for local production). Columns 1 and 3 are the shares of R&D by domestic and foreign firms, calibrated to match the data.

Columns 2 and 4 are inferred through the lens of the model. The second column is the share of R&D done by domestic firms at home for local production, measured by the revenue of the varieties developed, i.e., $\frac{\sum_d X_{oood}}{\sum_{m,d} X_{oomd}}$. These shares average 83.3%, reflecting that it is costly to separate production from

[5] Under the assumption that z^P is drawn from two Pareto distributions probabilistically, this integration is analytical:

$$\begin{aligned} \pi^R_{oi}(z^R) &= \int_0^\infty \pi^R_{oi}(z^P, z^R) g^P(z^P|z^R) dz^P \\ &= Prob(z^P_H|z^R) \cdot \int_0^\infty \pi^R_{oi}(z^P, z^R) dG_H(z^P) + Prob(z^P_L|z^R) \cdot \int_0^\infty \pi^R_{oi}(z^P, z^R) dG_L(z^P) \end{aligned}$$

Note from equation (B.11) that $\overline{\pi}_{oi}(z^P)$ is linear in a power function of z^P, and thus so is $\pi^R_{oi}(z^P, z^R) = (\gamma^{\frac{\gamma}{1-\gamma}} - \gamma^{\frac{1}{1-\gamma}})\left(\frac{1}{W^h_i}\right)^{\frac{\gamma}{1-\gamma}}\left(\overline{\pi}_{oi}(z^P) \cdot z^R\right)^{\frac{1}{1-\gamma}}$. With G^P_H and G^P_L being Pareto distributions, $\pi^R_{oi}(z^R)$ thus can be written as a function of z^R in an analytical form.

the headquarters and the R&D center at the same time. The fourth column reports the local production share of varieties developed by foreign affiliates. On average, 70% of foreign R&D leads to offshore production in the same host.[6] The shares tend to be higher in hosts with lower production costs.

Offshore R&D and the sources of national income. Offshore R&D enables firms to apply their knowhow globally and shape the sources of income for all. Columns 5 to 9 decompose the income of a country into manufacturing production, profit (total and that from overseas inventions), R&D, and marketing. In autarky, the sources of income are the same across countries. In the open economy, while marketing expenditures still account for a fixed share of income, the importance of other sources are altered by firms' global operations. In particular, advanced countries populated with the most efficient firms tend to earn a higher share of income from profit. For instance, profit accounts for 21% of income in the U.S., more than a third of which is from overseas R&D centers—an important part of the U.S. knowhow only realizes its value through offshore R&D. In contrast, only 5% of income in Slovakia is from profit and almost all of it is generated from varieties invented domestically. This finding shows that accounting for the returns through offshore R&D is important for valuing the intangible assets of nations.

SA.C.3 Implications for the Global Incidence of FDI and R&D Policies

Most existing quantitative research on MNCs does not differentiate policies on offshore R&D and offshore production. Yet policy makers usually have at their disposal instruments that specifically target each of these two activities. I examine whether not differentiating R&D and production is an important restriction for practical policy evaluations. As an example, I focus on two forms of FDI liberalization among emerging countries, which has gained significance in the past decade.

Integration among emerging countries. I first consider a reform eliminating the overhead cost for offshore R&D between a set of emerging countries, including Brazil, China, Hungary, Mexico, Poland, Russia, Romania, and Turkey. In practice, this reduction in cost can take the forms of speedier approval of entry, subsidized land, or tax credits for the upfront investment in R&D. The first column of Table SA.C.2 reports the results. Not surprisingly, this policy benefits emerging countries. Yet their benefits are at the expense of developed countries, whose overseas R&D centers in emerging countries have to face tougher competition after the policy.

The second experiment is liberalization in bilateral offshore production, which increases ϕ^P_{im} by 20% between the set of emerging countries. I focus on the different distributions of the welfare gains across countries, rather than the level of welfare gains, because these two types of liberalization do not necessarily have the same administrative burden or fiscal costs. As shown in Column 2 of Table SA.C.2, emerging countries still gain significantly, but differently from the first experiment, major developed countries are also better off—thanks to their presence in the emerging countries through offshore R&D, countries like the U.S. benefit from an increase in the profit of the varieties they develop there. These two experiments demonstrate that openness to R&D and production could have qualitatively different third country effects. This point is important for multilateral investment treaties, which often cover investment in intellectual properties.

Because of the within-firm linkages between trade and production, and offshore R&D, incorporating the latter also affects policies on trade and offshore production. To make this point, I consider the same liberalization as in the second experiment, but in a restricted version of the model without offshore R&D. The welfare impacts of this experiment are reported in the third column of Table SA.C.2. Compared to the baseline economy, emerging countries generally benefit less—without foreign entrants, the overall R&D in these host does not expand as much to take advantage of the increasing access to overseas producers. Developed countries experience net losses: the production of their affiliates in emerging

[6] That the majority of affiliate R&D is conducted for local production is consistent with Bilir and Morales (2020). Using a different data set (American MNCs) and a different approach (production function estimation), they find that R&D in an affiliate mostly applies to the affiliate itself and has limited spillovers on sibling affiliates.

Table SA.C.1: Decomposition of R&D and Sources of Firm Profit

	Source and use of R&D				Source of income (% of total income)				
	% by domestic firms		% by foreign firms		mfg.	profit		R&D	mkt.
Country		% of local prod.		% of local prod.		total	inventions abroad		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
AUS	60.2	87.8	39.8	75.5	82.7	6.4	0.2	2.6	8.3
AUT	54.5	79.3	45.5	61.4	82.1	6.8	0.4	2.7	8.3
BEL	41.1	64.0	58.9	47.2	85.0	4.6	0.7	2.0	8.3
BGR	80.8	97.3	19.2	87.8	79.2	9.3	0.0	3.2	8.3
BRA	42.5	97.5	57.5	84.8	76.6	10.6	0.0	4.5	8.3
CAN	47.5	85.5	52.5	68.9	75.5	11.9	2.4	4.3	8.3
CHE	58.1	71.8	41.9	51.6	74.2	13.1	2.2	4.3	8.3
CHN	57.8	95.2	42.2	71.9	76.6	10.2	0.0	4.8	8.3
CZE	73.2	93.0	26.8	81.4	83.8	5.7	0.0	2.2	8.3
DEU	71.5	71.1	28.5	60.6	76.9	11.2	1.7	3.6	8.3
DNK	61.6	67.2	38.4	45.9	79.2	9.7	2.6	2.8	8.3
ESP	77.9	91.1	22.1	80.0	81.4	7.6	0.1	2.7	8.3
EST	71.4	93.4	28.6	76.1	82.3	6.9	0.1	2.4	8.3
FIN	74.9	73.8	25.1	48.7	73.1	14.2	2.3	4.4	8.3
FRA	73.3	71.7	26.7	56.5	76.4	11.6	1.5	3.6	8.3
GBR	37.0	47.6	63.0	40.1	84.0	5.4	2.0	2.3	8.3
GRC	49.9	97.4	50.1	88.0	77.9	10.1	0.0	3.7	8.3
HRV	46.4	98.4	53.6	90.7	78.5	9.5	0.0	3.6	8.3
HUN	34.1	97.6	65.9	87.1	80.1	7.0	0.0	4.6	8.3
IRL	26.3	47.8	73.7	33.8	70.2	17.8	12.6	3.6	8.3
ITA	56.2	93.2	43.8	81.7	78.1	9.5	0.1	4.1	8.3
JPN	97.0	84.5	3.0	69.3	72.2	15.1	1.8	4.4	8.3
KOR	93.5	90.5	6.5	76.7	74.4	13.1	0.7	4.2	8.3
LTU	79.6	95.0	20.4	81.8	81.3	7.7	0.1	2.7	8.3
LVA	94.3	97.3	5.7	86.4	80.3	8.5	0.0	2.9	8.3
MEX	44.8	98.9	55.2	95.0	76.9	10.7	0.0	4.1	8.3
NLD	69.2	15.2	30.8	12.7	80.5	9.1	3.7	2.1	8.3
NOR	67.7	69.3	32.3	45.8	67.3	20.9	12.3	3.5	8.3
POL	85.3	92.4	14.7	81.6	81.7	7.5	0.1	2.5	8.3
PRT	73.7	95.0	26.3	83.7	81.3	7.7	0.0	2.7	8.3
ROU	45.5	98.6	54.5	94.6	84.3	5.1	0.0	2.3	8.3
RUS	88.5	92.0	11.5	82.4	79.1	9.4	0.1	3.2	8.3
SVK	72.0	94.7	28.0	83.4	85.1	4.8	0.0	1.7	8.3
SVN	82.9	94.9	17.1	80.1	79.7	8.9	0.0	3.1	8.3
SWE	56.7	75.6	43.3	55.0	78.0	9.7	0.8	4.0	8.3
TUR	79.8	98.0	20.2	89.8	75.5	12.0	0.0	4.2	8.3
USA	83.7	68.2	16.3	59.3	66.0	20.8	7.7	4.9	8.3
Mean	65.1	83.3	34.9	70.2	78.3	10.0	1.5	3.4	8.3

Notes: All numbers are in percent. Columns 1 and 3 report the source of R&D, i.e., whether it is by domestic firms (Column 1) or foreign firms (Column 3). Columns 2 and 4 report the fraction of R&D devoted to local production. Columns 5 to 9 decompose the fractions of income of a country from different activities: manufacturing production, profit (total and that accrued from products developed offshore), R&D, and marketing.

Table SA.C.2: Implications for FDI and R&D Policies

Country	FDI Integration Among Emerging Economies			Higher U.K. R&D Efficiency	
	off. R&D	off. Prod.	off. Prod. w/o off. R&D	Baseline	w/o off. R&D
	(1)	(2)	(3)	(4)	(5)
BRA	0.9	1.1	0.8	0.0	0.0
CHN	1.1	1.2	1.3	0.0	0.0
HUN	8.1	9.1	4.2	0.0	0.0
MEX	0.7	1.9	0.8	0.0	0.0
POL	1.6	7.9	3.9	0.1	0.0
ROU	10.0	19.8	10.3	0.1	0.1
RUS	2.5	4.7	2.7	0.0	0.0
TUR	1.0	1.5	0.9	0.0	0.0
DEU	-0.1	0.3	-0.1	0.0	0.0
FRA	-0.1	0.3	-0.1	0.0	0.0
GBR	0.0	0.3	-0.1	5.3	5.7
NLD	0.0	0.8	-0.2	0.4	0.2
JPN	-0.1	0.5	-0.2	-0.1	0.0
USA	-0.6	1.4	-0.3	0.1	-0.1
mean (all)	0.5	1.4	0.6	0.2	0.2

Notes: The first two columns show that offshore R&D and production policies between the same set of countries have qualitatively different incidences on the rest of countries. Comparison between Columns 2 and 3 shows that the same offshore production policy has different effects when offshore R&D is overlooked. The last two columns show that the spillover effects on the rest of the world of an increase in R&D efficiency in U.K. are different if offshore R&D is overlooked.

countries face an increase in competition as before, but now they cannot make up for the losses with the profit of the varieties they develop in these countries. The comparison between these two experiments shows that even if one's goal is solely to understand the effect of liberalizing offshore production, it is important to incorporate offshore R&D.

The global incidence of R&D policies. The presence of offshore R&D also implies that R&D policies can have a global impact simply because such policies would typically also apply to local R&D centers owned by foreign firms. This channel is independent of and in addition to the spillover effects of R&D across affiliates studied in the literature (e.g., Bilir and Morales, 2020). As an example, I consider a 20% increase in the efficiency of R&D taking place in the U.K.[7] This change increases the real income of the U.K. by 5.3%. Countries with extensive ties with the U.K. via offshore R&D are also better off. In total, 14% the total gains accrue to other countries. On the other hand, if offshore R&D is shut down, the same change in the R&D efficiency of the U.K. will benefit itself by 5.7%, which is 133% of the total gains—other countries, most notably the U.S., bear welfare losses.

References

Bilir, L Kamran and Eduardo Morales, "Innovation in the Global Firm," *Journal of Political Economy*, 2020, *128* (4), 1566–1625.

Lee, David L, Justin McCrary, Marcelo J Moreira, and Jack Porter, "Valid t-ratio Inference for IV," *arXiv preprint arXiv:2010.05058*, 2020.

OECD, "The Analytical AMNE Database," http://stats.oecd.org/Index.aspx?DataSetCode=AFA_CALC_IN3, 2015. Accessed: 2015-11-24.

[7]For foreign firms, this change resembles the 'patent box,' a policy implemented in the U.K. to attract innovative firms that reduces the corporate tax rate on revenues generated through R&D.

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

_ , _ , **and** _ , "Data For 'Multinational Production: Data And Stylized Facts'," ISPCR, https://doi.org/10.3886/E113384V1, 2019.