Online Appendix for the Paper "A Global Analysis of Factors Impacting the Intensive and Extensive Margins of Bilateral FDI"

Nguyen, A.T.N (2019), A Global Analysis of Factors Impacting the Intensive and Extensive Margins of Bilateral FDI, World Economy, 42(9), p.2649–2667.

As emphasised in Gujarati and Porter (2009, p. 30) "the researcher should always keep in mind that the results of research are only as good as the quality of the data". Data series that particularly require more assessment are bilateral FDI, bilateral investment treaties (BITs), common religion and common legal origin. Data and Stata codes are available on request.[1]

1. Country List

In the scope of this research, developed countries (DCs) are countries in the high income bracket according to the World Bank's 2008 data on global income groups. The rest of the world is considered less developed countries (LDCs).

43 DCs: Australia, Austria, Bahamas, Bahrain, Barbados, Belgium, Canada, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hong Kong, Hungary, Iceland, Ireland, Israel, Italy, Japan, South Korea, Kuwait, Luxembourg, Malta, Netherlands, New Zealand, Norway, Oman, Portugal, Qatar, Saudi Arabia, Singapore, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Trinidad and Tobago, United Kingdom, United States.

67 LDCs: Argentina, Armenia, Azerbaijan, Belarus, Benin, Bolivia, Bosnia & Herzegovina, Botswana, Brazil, Bulgaria, Burkina Faso, Cameroon, Chile, China, Colombia, Costa Rica, Cote d'Ivoire, Dominican Republic, Ecuador, Egypt, Fiji, Gabon, Georgia, Guatemala, Guinea, Honduras, India, Indonesia, Iran, Iraq, Jamaica, Jordan, Kazakhstan, Kenya, Kyrgyz Republic, Latvia, Lebanon, Lithuania, Malaysia, Mauritius, Mexico, Moldova, Mongolia, Morocco, Mozambique, Namibia, Nicaragua, Nigeria, Panama, Paraguay, Peru, Philippines, Poland, Russian Federation, Rwanda, Senegal, South Africa, Sri Lanka, Swaziland, Tanzania, Thailand, Tunisia, Turkey, Ukraine, Uruguay, Venezuela, Zimbabwe.

2. Bilateral FDI Data

2.1. Overview of the Database

[1] Email: Anh T.N. Nguyen, ngocanhnguyen2312@gmail.com

The two most popular databases for time-series bilateral FDI are UNCTAD (United Nations Conference on Trade and Development) and the OECD (Organisation for Economic Co-operation and Development). While UNCTAD provides bilateral FDI data for 206 countries in the world, the OECD reports data only for its 35 members. Therefore, the OECD database does not capture FDI between non-member countries but this is still a useful source to compare with UNCTAD and to assess data quality in general. UNCTAD has the biggest database for bilateral FDI and has been employed in various studies such as Bergstrand and Egger (2007), Baccini and Dür (2013), and Dixon and Haslam (2016). Each country comes with four data series: **inflows, outflows, instock**, and **outstock.** Since UNCTAD updates its FDI data yearly, the following comments are for the data downloaded on 1st May, 2017 from the online database.[2]

Firstly, data from UNCTAD are inconsistent. In the UNCTAD database, if the country does not report data for one or more series, UNCTAD uses mirror data from partner countries or other international organizations. Obviously, we would expect that the mirror data are exactly the same as the original data. Surprisingly, this is not always the case. For instance, for the first ten years, 2001-2010, Singapore's mirror outflows to Korea exactly equal Korea' original inflows from Singapore. Nevertheless, Korea reports receiving 492 and 963 million USD from Singapore in 2011 and 2012, respectively (inflows data series of Korea) while Singapore's mirror outflows to Korea in these years are 0 (outflows data series of Singapore). Table A1 lists other examples of the inconsistent "mirror – origin" country pairs. The total number of inconsistencies in the database is 363 pairs.

Table A1: Examples of Inconsistent Mirror-Origin FDI Flows (million USD)

Year	Host	Source	Data type	Host country's Inward data		Source country's Outward data	
				Value	Data source	Value	Data source
2005	DZA	EGY	Flows	108.2	origin	0	mirror
2006	DZA	EGY	Flows	157.1	origin	0	mirror
2006	MMR	SGP	Flows	160.8	origin	0	mirror
2007	DZA	EGY	Flows	102.3	origin	0	mirror
2012	CYM	KOR	Stock	0	mirror	2788.9	origin
2012	MHL	KOR	Stock	0	mirror	540.1	origin
2012	MMR	KOR	Stock	0	mirror	1326.8	origin
2012	VNM	KOR	Stock	0	mirror	7181.7	origin

[2] http://unctad.org/en/Pages/DIAE/FDI%20Statistics/FDI-Statistics-Bilateral.aspx , accessed on 1st May, 2017.

In addition, although UNCTAD generally reports country pairs that do have an FDI relationship in at least one year over the period (2001 – 2012), in some countries' data series, they also include in the list of partner countries several economies that have absolute zero or negligible value reported as zero FDI over the period. For example, Georgia is in the list of Denmark's inflows series but all inflows to Denmark from Georgia are zero over the period. Technically, these zeros are not different from no FDI to other countries outside the list and therefore I do not include these observations in counting the number of reported observations in Table A3.

Lastly, in both UNCTAD and OECD databases, some countries report non-zero FDI sent to themselves such as Japan in OECD and Austria in UNCTAD. There is no explanation for these data and I exclude them in my analyses.

2.2. Inconsistency Between Data Series

In principle, FDI inflows to country *i* from country *j*, reported by country *i*, should equal FDI outflows from country *j* to country *i*, reported by country *j* [i.e. presuming they are all measured in the same currency and converted consistently]. And, the sum of all FDI inflows should match the sum of all FDI outflows. Similarly for inward and outward stocks. Unfortunately, because of the variations in interpretation of the definition of FDI and in the data compiling methodologies, inbound and outbound FDI data rarely match (Fujita, 2008; Gouel et al., 2012). This is similar to trade data. However, statistic publications give a very general picture of the discrepancies in total global inward and outward FDI but do not scrutinise the problem in detail.

Data discrepancies in bilateral FDI data have not been investigated or well-documented in any empirical studies. Empirical researchers may or may not be aware of this problem and simply use one data series for flows (inflows or outflows) and/or one series for stocks (instock or outstock) without justifying their choice. Therefore, this section aims to fill this gap, and give researchers a clearer idea about the choice of bilateral FDI data, i.e. what data series to use: flows or stocks, inbound or outbound, and why?

a. Stocks or Flows?

Table A2 demonstrates the correlations and maximum differences in data series from UNCTAD and OECD. The number of countries in UNCTAD database decreases to 201 after

I exclude all mirror data and the countries that do not report any data series originally. In addition, although there are more than 200 economies in the whole OECD database, the number of countries entering the correlation is only 35 which are all the reporters of OECD data, i.e. 35 OECD members. This is because only the country pairs with both in the OECD report data from both directions (inbound and outbound).

As demonstrated in Table A2, the discrepancies in FDI data series from either OECD or UNCTAD are almost the same although UNCTAD does have considerably more developing countries originally reporting data in their database. In particular, between outflows and inflows data series the correlation is less than 0.5 and the maximum difference is more than 100 million USD in both data sources. If we looked only at the data from UNCTAD, we would think that the discrepancies come from low quality data reported by developing countries compared to the same data reported by developed countries. But, Table A2 shows that data discrepancies in bilateral FDI are a problem in both databases. Moreover, the majority of the pairs that have large differences actually come from developed country pairs. For instance, Belgium reports inflows from Luxemburg in 2008 at around 63.8 billion USD while Luxemburg reports outflows to Belgium in the same year at approximately -48 billion USD in both databases. Moreover, though not shown in the table, the correlation between the same data series reported by UNCTAD and OECD is very high, i.e. above 0.97 for flow series and above 0.99 for stock series. Thus, discrepancies in bilateral FDI data are a problem for both developed and developing countries.

Table A2: Correlations and Differences between Data Series

		UNCTAD	OECD
Between inflows and outflows	Correlations	0.475	0.409
	Maximum difference	111.854 billion USD	110.819 billion USD
Between instock and outstock	Correlations	0.795	0.775
	Maximum difference	574.718 billion USD	562.902 billion USD
Coverage of countries		201 economies	35 OECD economies

However, according to Table A2, using data on FDI stocks is likely to give more reliable and consistent results than FDI flows as the bilateral stock data reported by different countries are much more highly correlated than flow data. Also, FDI stocks are considered to be more stable with less fluctuation than FDI flows. These findings agrees with Wacker (2016) who also shows that FDI stock is a better proxy for measuring multinational firms' activities.

On the other hand, FDI flows still possess some important advantages. First, values of FDI flows are more current and matching with other yearly data series, especially those in monetary units such as annual GDPs and wages. According to UNCTAD, many countries report FDI stocks by the cumulative sum of FDI flows over a number of years. FDI stocks this year may be calculated by simply adding current FDI flows to FDI stock value last year. The past value of FDI stocks is not adjusted for inflation and thus the FDI stock value this year does not reflect the current value of the investment made a few years ago. Additionally, FDI flow data can provide more accurate information for the yearly binary response of whether or not to invest as there are cases where the value of FDI stocks can be positive but the amount of FDI flows in the same year is actually zero. Simply speaking, both data series have their own advantages and disadvantages and research should employ both data series to draw concrete conclusions.

b. Negative FDI

More complicated than trade data, bilateral FDI data do not only contain excessive zeros but also have negative values. In this study, I ignore negative FDI as these observations are included in the participation equation only and considered as "no investment" outcome. Negative FDI stocks are generally the consequences of calculating methods (Gouel et al., 2012), i.e. FDI stocks may equal the accumulation of negative FDI flows over a period of time, and these negative values do not have any real economic meaning. Meanwhile, According to UNCTAD, negative FDI flows are explained as disinvestment and they have a real economic meaning. In the scope of this study, I do not examine disinvestment.

c. Inward or Outward?

Inward/inbound data series (inflows and instock) are reported by the host/receiving country while outward/outbound data series (outflows and outstock) are reported by the source/parent country. Generally speaking, inbound data reported by the receiving country are usually recorded with more accuracy because foreign investors have to go through a number of registration processes and pay local taxes and fees when investing in that country. Therefore, the host country government can keep track of the flows or stocks going to their country and the MNE's activities to some extent. Meanwhile, outbound data are more difficult to track from the parent country's point of view as the investment activity happens abroad. This is similar to the case of bilateral trade where import data are more accurate than export data

because imports create tariff revenues in the destination country while exports do not. Inbound data series are not only more accurate but also provide more observations than outbound data series. This is not surprising as countries tend to report more data on the series they can observe with less difficulty. This is shown in Table A3. The number of observations of inbound data series is always higher than that of outbound data series for both stocks and flows in the UNCTAD database.

Table A3: Descriptive Statistics on FDI data from UNCTAD Database (million USD)

	Series	Obs	Mean	Std. Dev.	Min	Max
Flows	Inflows	58583	235.9	2290.4	-59599.4	117617.9
	Outflows	40312	350.1	2859.7	-51212.0	109097.0
Stocks	Instock	52279	2661.1	16694.15	-12850.9	592273.2
	Outstock	41452	3633.8	20836.49	-30237.1	645098.0

Nevertheless, relying only on one data series (either inbound or outbound) as most previous empirical papers have done is also problematic. For example, United Arab Emirates (ARE) originally reports inflows from a total of 10 countries (ARE's inflows series). However, according to outflows series reported by other countries, apart from that list of 10 countries, there are another 25 economies that actually have FDI flows sent to ARE such as Belgium, Switzerland, China, Cyprus, Korea, etc. over the period 2001-2012. Therefore, using only one data series and assuming that any country not in the list has no FDI with that reporting country is incorrect.

The above imperfections in bilateral FDI encourage me to create more comprehensive data series on bilateral FDI from the UNCTAD database. I first remove all the mirror values reported by UNTACD and create a variable *stock* equalling inward stock series. This mean I give the first priority to inbound data. Then, it is replaced by the outward stock series if inward stock is missing or non-positive. Afterwards, in any year when the data are missing or negative, according to reports from both countries in the pair, it is considered no FDI. This means there are no missing data for the dependent variable in this study. Other studies that use data from UNCTAD and claim that they rely on only one series, either inward or outward, without further clarification, actually partly follow this practice. This is because the original data from UNCTAD may include mirror data in some countries when reporters do not report that series in their database. A similar process is applied to create a variable *flow* with the highest number of positive dyadic observations.

3. Bilateral Investment Treaty

UNCTAD provides data on BITs between 233 economies. Empirical studies employing data on BITs often assume that the BITs are forever into force after the year of entry into force. Thus, the dummy for BITs is set to equal 1 for all the years after the BITs come into force. However, inspecting BIT data shows that this is not the case.

Table A4: Descriptive Statistics on BITs from UNCTAD Database[3]

Time span		1962-2017
Number of economies		233
Number of BITs		3220
Number of terminated BITs by reasons of termination	Replaced by new treaty	130
	Unilaterally denounced	64
	Terminated by consent	19
	Expired	3
	Total	216

Table A4 shows that nearly 7% of BITs are terminated. However, more than half of these terminated BITs are replaced by a new treaty coming into force in the same year as the year of termination. This means in these cases the dummy variable for BITs for that pair of country still takes the value of 1 and is not affected by the termination. On the other hand, approximately 3% of BITs are terminated without a consecutive new treaty coming into force and leads to wrong values of variables related to BITs. In my dataset, the dummy for BITs receives the value of 1 between the year of entry into force to the year before (if relevant) the termination year.

4. Common Religion and Legal Origin

There are two variables that I construct, common legal origin and common religion index, following the ideas of Helpman et al. (2008), Santos Silva et al. (2014), and Garrett (2016). These studies construct a common religion index by multiplying the percentages of the population of source and host countries that share one of the three main religions - Muslim, Protestant, and Catholic. As this index is originally proposed by Helpman et al. (2008), I will refer to as the Helpman religion index. This index leads to misleading results in many cases, especially for Asian countries. For example, according to the Helpman index, Japan and Thailand's common religion index is absolute zero as the two countries do not share any of the above three religions. However, Japan has 66.8% of its population following Buddhism

[3] http://investmentpolicyhub.unctad.org/IIA/IiasByCountry#iiaInnerMenu, accessed on 2nd May, 2017.

and the number for Thailand is 94.6%. If Buddhism was included in the Helpman index, the common religion index for the two countries would be 63.19. Therefore, for a more accurate index, the common religion index in this study equals to:

$$\frac{\sum_k \%\ \text{religion } k \text{ in source country x } \%\ \text{religion } k \text{ in host country}}{100} \qquad \text{(A1)}$$

(A1) captures all possible common religion between each pair of countries, which is the most accurate way to capture the share of common religions between two countries. But, it does not come without costs. Text data on religion from the CIA World Factbook[4] are in fact very messy and lack precise information. First, in many cases they provide the total percentage of the population following a number of religions but not each separate religion. In these cases, I manually collect and update the data for each religion in some countries from alternative data sources, such as Wikipedia and countries' statistics. It is important that the manually imported information comes with clear and reliable references. If such data are not available, I equally divide the total percentage to each religion in the aggregate list. This process is time-consuming and may explain why the other studies follow the more restrictive Helpman religion index. However, data for the three religions in the Helpman index suffer from the same problem as other religions. For example, many countries report data only on total percentages of people following Christianity (Tazania 61.4%, Tongo 29%, Nigeria 40%, and so on) without clarifying whether they are Protestant or Catholic. Hence, I also manually check other sources for these unclear values. Due to the imprecision in religion data, the level of accuracy of this index has potential for future improvement.

Lastly, data to construct the dummy for a common legal origin are also collected from the CIA World Factbook.[5] According to the CIA, the legal systems of nearly all countries are generally modelled upon elements of five main types: civil law; common law; customary law; mixed or pluralistic law; and religious law. Among these, mixed or pluralistic law is just a combination of the other four systems. Therefore, the common legal origin is a dummy variable taking on a value of 1 if the two countries share at least one of the four roots in their legal system: civil law, common law, customary law and religious law.

[4] https://www.cia.gov/library/publications/the-world-factbook/fields/2122.html, accessed on 30th May, 2017.
[5] https://www.cia.gov/library/publications/the-world-factbook/fields/2100.html, accessed on 20th, June, 2017.

Table A5: Intensive Margin – Country Groups – FDI Stock

	Expected Sign	DCs to DCs		DCs to LDCs		LDCs to DCs		LDCs to LDCs	
		FE	CMRE	FE	CMRE	FE	CMRE	FE	CMRE
lgdp_s	+	0.614***	0.594***	0.515***	0.471***	0.200	0.176	-0.005	-0.054
		(0.145)	(0.144)	(0.152)	(0.150)	(0.214)	(0.210)	(0.193)	(0.190)
lgdp_h	+	0.555***	0.561***	0.288**	0.304***	0.615*	0.547*	0.463***	0.453***
		(0.144)	(0.144)	(0.116)	(0.115)	(0.336)	(0.327)	(0.165)	(0.162)
tech_s	+	2.653***	2.659***	1.042	1.123	4.090***	4.072***	2.523***	2.682***
		(0.996)	(0.995)	(1.089)	(1.084)	(1.116)	(1.107)	(0.919)	(0.914)
delta_h	-	3.296	2.196	7.477	6.722	18.300	18.973	-12.453	-10.084
		(9.256)	(9.202)	(7.584)	(7.504)	(21.331)	(20.752)	(11.555)	(11.270)
remote_s	-	-1.992	-2.041	-1.391	-1.572	8.727***	7.863***	2.256	1.708
		(1.734)	(1.730)	(1.605)	(1.593)	(3.050)	(3.012)	(3.362)	(3.274)
remote_h	-	1.921	1.904	5.422***	5.433***	-3.098	-2.521	-0.786	-0.448
		(1.214)	(1.212)	(1.686)	(1.677)	(3.348)	(3.285)	(3.093)	(3.031)
startup_h	-	-0.083***	-0.086***	0.022	0.028	0.084	0.087	-0.021	-0.017
		(0.025)	(0.025)	(0.028)	(0.028)	(0.056)	(0.056)	(0.045)	(0.044)
phi_h	+	5.676***	5.232***	3.090***	3.124***	5.510	6.403**	1.642***	1.486**
		(1.229)	(1.168)	(0.364)	(0.359)	(3.455)	(2.887)	(0.597)	(0.584)
BIT	+	-0.356	-0.363	0.238**	0.208**	-0.068	-0.101	0.096	0.115
		(0.275)	(0.273)	(0.106)	(0.105)	(0.203)	(0.198)	(0.177)	(0.172)
PTA	+	0.105	0.108	0.173**	0.174**	0.129	0.130	-0.031	-0.013
		(0.120)	(0.120)	(0.079)	(0.079)	(0.174)	(0.171)	(0.149)	(0.147)
currency	+	0.246*	0.241*		1.179***		1.823		0.058
		(0.126)	(0.126)		(0.387)		(1.155)		(0.370)
ldist	-		-0.678***		-0.797***		-0.840***		-0.721***
			(0.115)		(0.101)		(0.124)		(0.117)
border	+		0.242		0.585		0.735		0.693***
			(0.270)		(0.496)		(0.662)		(0.246)
language	+		1.023***		0.567***		0.895***		0.586***
			(0.213)		(0.161)		(0.235)		(0.183)
coloniser	+		0.561		0.376		1.186***		-0.047
			(0.501)		(0.343)		(0.453)		(0.210)
colony	+		0.913***		1.625***		1.557***		1.009***
			(0.297)		(0.258)		(0.410)		(0.388)
legal	+		0.148		-0.297**		-0.023		0.117
			(0.149)		(0.119)		(0.164)		(0.183)
religion	+		0.094**		0.115***		0.067*		0.026
			(0.042)		(0.029)		(0.035)		(0.028)
Observations		9203	9203	9468	9468	4492	4492	4788	4788

Standard errors in parentheses: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A6: Intensive Margin – Country Groups – FDI Flows

	Expected	DCs to DCs		DCs to LDCs		LDCs to DCs		LDCs to LDCs	
	Sign	FE	CMRE	FE	CMRE	FE	CMRE	FE	CMRE
lgdp_s	+	0.919***	0.881***	0.977***	0.942***	0.155	0.227	0.109	-0.046
		(0.231)	(0.226)	(0.207)	(0.205)	(0.298)	(0.283)	(0.269)	(0.263)
lgdp_h	+	0.636**	0.669***	0.780***	0.784***	0.372	0.132	0.610***	0.655***
		(0.249)	(0.244)	(0.132)	(0.131)	(0.378)	(0.365)	(0.217)	(0.212)
tech_s	+	-0.250	-0.040	0.515	0.387	4.165***	3.902***	3.125***	3.562***
		(1.575)	(1.560)	(1.640)	(1.620)	(1.185)	(1.164)	(1.013)	(0.992)
delta_h	-	27.013*	24.689*	23.193**	23.042**	57.469**	53.838**	-7.352	-3.838
		(14.504)	(14.016)	(9.222)	(9.024)	(24.440)	(23.760)	(13.086)	(12.495)
remote_s	-	0.029	-0.047	0.900	1.236	-3.042	-2.494	1.878	0.058
		(2.122)	(2.088)	(1.849)	(1.823)	(3.818)	(3.623)	(3.683)	(3.454)
remote_h	-	-2.632	-2.772	-6.921***	-6.719***	-0.863	-0.765	-15.458***	-12.997***
		(1.933)	(1.915)	(2.152)	(2.113)	(3.779)	(3.579)	(3.874)	(3.687)
startup_h	-	-0.144***	-0.137***	-0.056*	-0.047	0.020	0.016	0.010	0.022
		(0.039)	(0.039)	(0.032)	(0.032)	(0.067)	(0.065)	(0.052)	(0.050)
phi_h	+	4.839***	3.165**	1.422*	1.382*	1.781	3.479	2.365**	2.257*
		(1.771)	(1.589)	(0.767)	(0.762)	(3.140)	(2.509)	(1.117)	(1.183)
BIT	+	-0.342	-0.361	0.051	0.030	0.492*	0.417	0.118	0.147
		(0.245)	(0.244)	(0.133)	(0.132)	(0.277)	(0.262)	(0.123)	(0.123)
PTA	+	0.357**	0.355**	0.024	0.034	0.125	0.026	-0.066	-0.086
		(0.149)	(0.151)	(0.100)	(0.099)	(0.192)	(0.183)	(0.156)	(0.153)
currency	+	0.306**	0.305**		0.498		1.601		-0.121
		(0.156)	(0.149)		(0.373)		(1.033)		(1.251)
ldist	-		-0.590***		-0.734***		-0.682***		-0.510***
			(0.098)		(0.085)		(0.100)		(0.097)
border	+		0.246		0.358		0.472		0.430**
			(0.220)		(0.485)		(0.481)		(0.205)
language	+		0.850***		0.231		0.881***		0.377**
			(0.173)		(0.171)		(0.221)		(0.178)
coloniser	+		0.306		0.321		1.139***		0.404*
			(0.524)		(0.444)		(0.317)		(0.216)
colony	+		0.822***		1.228***		0.951***		1.089***
			(0.249)		(0.248)		(0.330)		(0.297)
legal	+		-0.061		-0.356***		-0.039		0.067
			(0.127)		(0.103)		(0.132)		(0.160)
religion	+		0.048		0.104***		0.063**		0.022
			(0.037)		(0.025)		(0.028)		(0.023)
Observations		6971	6971	7540	7540	3195	3195	3487	3487

Standard errors in parentheses: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

REFERENCES

Baccini, L., & Dür, A. (2013). Investment Discrimination and the Proliferation of Preferential Trade Agreements. *Journal of Conflict Resolution, 59*(4), 617-644.

Bergstrand, J. H., & Egger, P. (2007). A Knowledge-and-Physical-Capital Model of International Trade Flows, Foreign Direct Investment, and Multinational Enterprises. *Journal of International Economics, 73*(2), 278-308.

Dixon, J., & Haslam, P. A. (2016). Does the Quality of Investment Protection Affect FDI Flows to Developing Countries? Evidence from Latin America. *The World Economy, 39*(8), 1080-1108.

Fujita, M. (2008). A critical assessment of FDI data and policy implications. *Transnational Corporations, 17*(2), 107-126.

Garrett, J. Z. (2016). Explaining Asymmetries in Bilateral FDI Flows. *International Review of Economics & Finance, 41*(2016), 155-171.

Gouel, C., Guimbard, H., & Laborde, D. (2012). A Foreign Direct Investment database for global CGE models. Working Paper 2012. Paris: CEPII.

Gujarati, D. N., & Porter, D. C. (2009). *Basic Econometrics* (5 ed.). New York: McGraw-Hill Irwin.

Helpman, E., Melitz, M., & Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics, 123*(2), 441-487.

Santos Silva, J. M. C., Tenreyro, S., & Wei, K. (2014). Estimating the extensive margin of trade. *Journal of International Economics, 93*(1), 67-75.

Wacker, K. M. (2016). (When) Should We Use Foreign Direct Investment Data to Measure the Activities of Multinational Corporations? Theory and Evidence. *Review of International Economics, 24*(5), 980-999.