

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<https://hdl.handle.net/2066/219782>

Please be advised that this information was generated on 2021-03-03 and may be subject to change.



Cross-national distance as an explanatory variable in international management: Fundamental challenge and solution



André van Hoorn

Radboud University, Institute for Management Research, PO Box 9108, 6500 HK Nijmegen, the Netherlands

ARTICLE INFO

Keywords:

Distance metrics
Location-identification problem
Omitted variables
Distances
Pure-distance approach

ABSTRACT

Although recent work provides insightful theoretical and practical suggestions for improving contextual distance research in international management, the fundamental problem with using distance indicators as explanatory variables remains too little recognized and largely unaddressed. This problem is that cross-national distance metrics partially identify host and/or home countries in one's sample, what I term location-identification. Location-identification can occur irrespective of the number of home/host countries considered and means that a distance indicator partly captures country fixed effects when used as an independent or explanatory variable. As a result, in empirical distance research, genuine distance effects are often confounded with country-specific measurement error in the dependent variable as well as with direct effects due to various home- or host-country features. I present empirical evidence on the pervasiveness of this critical challenge to cross-national distance research and propose a practical and effective solution for addressing it, which is to use "pure" distance indicators that are cleansed from confounding home- and host-location influences.

1. Introduction

The concept of home-host country or cross-national distance has been broadly adopted as one of the quintessential constructs in international management (Ghemawat, 2016; Hutzschenreuter et al., 2016; López-Duarte et al., 2019; Verbeke et al., 2017; Yeganeh, 2014). At the same time, researchers in the field have long been concerned with various limitations of the cross-national distance construct, particularly of cultural and institutional distance (e.g., Ambos and Håkanson, 2014; Bae and Salomon, 2010; Shenkar, 2001; Tung and Verbeke, 2010; Zaheer et al., 2012). Concerns raised, include, among others, the possibility of asymmetric distance effects (Chikhouni et al., 2017; Huang et al., 2017; Konara and Shirodkar, 2018) and the neglect of intra-country diversity in common distance indicators (Beugelsdijk et al., 2015).

More generally, a major issue in the distance literature is whether estimated distance effects reflect genuine distance effects or are perhaps driven by other, country-level factors, the problem of omitted variables (e.g., Beugelsdijk et al., 2017; Beugelsdijk et al., 2018a, 2018b; Harzing, 2003; Kirkman et al., 2006). On this latter question, there is mounting theoretical and empirical evidence that observed effects of cultural or institutional distance may often reflect effects directly due to countries' cultural or institutional profiles. Because of a lack of diversity in the home and host countries considered, in many distance studies, indicators of cultural or institutional distance correlate rather strongly with home or host countries' cultural and institutional profiles (Harzing and Pudelko, 2016; Van Hoorn and Maseland, 2014, 2016). Such distance-profile correlations can, in turn, create a confounded variables problem (Brouthers et al., 2016). Specifically, it becomes difficult to distinguish effects genuinely due to distance from direct effects due to

E-mail address: A.van.Hoorn@fm.ru.nl.

<https://doi.org/10.1016/j.intman.2020.100773>

Received 25 November 2019; Received in revised form 17 April 2020; Accepted 19 June 2020

1075-4253/© 2020 The Author. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

home or host countries' cultural or institutional profiles. Practically, it seems that many distance studies, particularly studies involving samples comprising only one home or one host country, have been conflating effects due to countries' cultural or institutional profile with effects due to cultural or institutional distance. This has been called the problem of distance-profile conflation (Van Hoorn and Maseland, 2016). Finally, Kostova et al. (2019, p. 23) have recently raised a similar but more general question concerning the institutional distance literature, namely whether found effects “are [...] really due to institutional distance or some undefined generic variable that captures country?”

This paper extends the above literature, which focuses on a subset of cross-national distance indicators, to consider an issue that affects different forms of cross-national distance as explanatory or independent variables. Specifically, this paper argues theoretically and demonstrates empirically that cross-national distance indicators can partially identify host and/or home locations in a sample, what can be called location-identification. Implication is that cross-national distance indicators may partly capture host and/or home country fixed effects when used as explanatory variables, what can be called the location-identification problem. This problem is pervasive for four main reasons. The first is that distance metrics' ability to capture country fixed effects implies a rather high risk of upwardly or downwardly biased estimates of the impact of measured distance on a given dependent variable. The rationale is that country fixed effects are nearly always significant, not least because these fixed effects are able to pick up any measurement error in the dependent variable that is specific to a country in one's sample.¹ The second is that location-identification is not limited to samples with relatively few home or host countries. This is again because the ability to identify specific locations is driven by the presence of country-specific variation in measured cross-national distance, which can exist in any sample, irrespective of its size. The third is that distance metrics' ability to capture country fixed effects does not only bias estimated distance effects but also results of any empirical analysis that includes cross-national distance as an independent variable. Specifically, there are many studies in international management that focus on country-level factors as explanatory variables and only use cross-national distance as a control variable (cf. Note 1). However, when distance is partly capturing country fixed effects, adding distance as a control variable will also introduce upward or downward biases in the estimated effects of such country-level factors. The fourth is that location-identification is driven by the presence of country-specific variation in measured home-host country distance. Hence, as indicated, location-identification is not limited to distance metrics involving differences between countries but could occur when using other types of cross-national distance indicators as well.

This paper is organized as follows. In the next section, I briefly discuss the theoretical mechanism underlying cross-national distance indicators' ability to identify particular home/host locations. Section 3 presents empirical evidence. I create two extreme hypothetical benchmarks that enable me to illustrate empirically how location-identification is a theoretical issue that can affect different types of cross-national distance indicators. A critical question for cross-national distance research is how to overcome the statistical biases that appear inherent to the use of cross-national distance metrics as independent variables. In Section 4 I propose and empirically illustrate an approach to addressing the confounded variables problem that can be caused by location-identification. My proposed approach is both practical and rigorous and involves using so-called “pure” distance indicators that are cleansed from confounding home- and host-location influences. I end with a discussion of the location-identification problem in relation to distance-profile correlations and the resulting problem of distance-profile conflation (e.g., Van Hoorn and Maseland, 2016). Although the two problems are related, the problem of location-identification is much broader and pervasive and some of the insightful means for improving empirical distance research (e.g., Beugelsdijk et al., 2018a) are insufficient for successfully addressing the location-identification problem.

2. Outlier countries, country-specific distances, and the partial identification of location

Location-identification occurs when a country is an outlier on a distance metric and therefore starts to become uniquely identified by the measured distance associated with it. In some samples, measured distance between home and host countries is almost entirely random and there is no systematic relationship between the locations of the home and host countries and the measured distance “to” or “from” these countries (see Panel A of Fig. 1 for a graphical illustration involving four home countries [M] and four host countries [S]). Many other samples, however, comprise outlier countries and as a result measured distance in these samples is partly specific to the home and host countries considered. A home (host) country is an outlier country when, on average, it is either much farther from or much closer to the host (home) countries in the sample than most other home (host) countries are. Panel B of Fig. 1 provides a graphical illustration involving four home countries, one of which, country M_F (F for far), is an outlier country.

The ramification of country-specific variation in measured distance is that distance indicators have the ability to provide partial identification of one or more countries in a sample. The reason is that when a country is a relative outlier, it distinguishes itself from other countries in the sample by the average amount of measured distance to and from this outlier country. Specifically, outlier countries are characterized either by having relatively large home-host country distances or by having relatively small home-host country distances. In fact, as also illustrated by Fig. 1, when there are outlier countries, some measured home-host country distances, either large ones or small ones, can only come from one or a few home or host countries. To wit, in Panel B, all cases in which measured distance is above the average home-host country distance in the sample involve country M_F as the home country.

The general implication is that, observing certain amounts of home-host country distance, it becomes possible to partially identify (or accurately predict) some of the home and/or host countries involved. In real-world samples it is therefore likely that measured

¹ In addition to plain country-specific measurement error in the dependent variable, there are of course many substantive ways in which home or host countries matter for MNEs (e.g., Estrin et al., 2017; Geleilate et al., 2016).

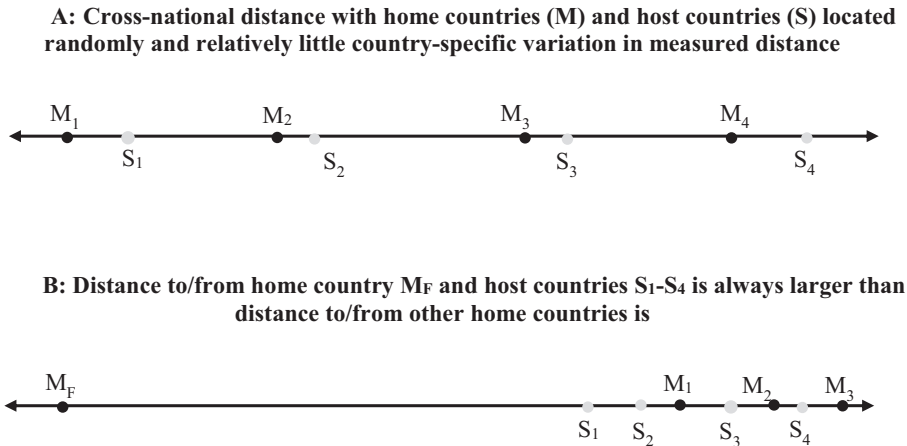


Fig. 1. Cross-national distance with relatively little or relatively much country-specific variation in measured home-host country distance
Notes: Because home country M_F (F for far) in Panel B is, on average, much farther from the host countries (S_1 - S_4) in the sample than the other home countries (M_1 - M_3) are, large home-host country distances always involve M_F as the home country.

distance partly captures home and/or host country fixed effects when considered as an explanatory variable. Moreover, this problem is independent from the number of home and/or host countries in one's sample. Location-identification does not occur because samples are too small. Instead, location-identification occurs because there is country-specific variation in measured home-host country distance, which can exist in samples of any size.

3. Empirical evidence on location-identification

3.1. Empirical approach

Theoretically, the location-identification problem is that cross-national distance indicators partially identify home/host locations, which, in turn, means that these indicators are prone to rendering biased empirical results. However, empirically, a key open question is exactly how much information on countries can be captured by distance indicators. In this section, I analyze two different samples of home-host country dyads using multinomial logistic regression models (Long and Freese, 2014). In these models, location, that is the different home/host countries in a sample, is the categorical dependent variable whereas measured distance is the independent variable. If measured distance contains a substantial amount of information on home or host locations, measured distance will be a statistically significant predictor of one or more of the host and/or home countries that are included in a sample. For each multinomial model, I consider the statistical significance (and overall fit) of the model with measured home-host country distance as the only predictor variable. However, I also estimate the marginal probabilities for each country/location in the sample. For all analyses, I follow the guideline set by Van Hoorn and Maseland (2016) of considering minimum seven home and minimum seven host countries.

3.2. Samples

As mentioned, for my empirical evidence I consider two hypothetical benchmarks. The degree to which location-identification occurs in the real world depends on many factors. As a result, it is entirely possible that a researcher ends up studying a specific sample and a specific cross-national distance metric for which the location-identification problem is marginal. The two extreme benchmarks help establish that the phenomenon of location-identification is a theoretical possibility but not a certainty. For both benchmarks, I create two hypothetical samples comprising 10 home and 10 host countries.

For the first sample (Sample 1), I aimed to create a sample that comprises a comparatively large amount of variation in measured cross-national distance that is specific to the home and the host countries in the sample. This sample comprises 20 different countries and 100 ($=10 \times 10$) home-host country dyads (Table 1). For the second sample (Sample 2), I aimed to create a sample that comprises a comparatively large amount of variation in measured cross-national distance that is specific to the home and the host countries in the sample. In this sample, home and host countries are the same. Hence, this sample comprises 10 different countries and 90 ($=10 \times 10-10$) home-host country dyads (Table 2).

To be sure, I have constructed these two samples without reference to any underlying profile score or to countries' exact locations in a given space. However, for the sake of completeness, Table A1 in Appendix A presents details on possible locations of the countries in these two samples that match the distances presented in Tables 1 and 2. Note, though, that these locations have no quality trait associated with them and are therefore irrelevant, only measured home-host country distance matters.

Table 1
Distances between the countries in benchmark Sample 1.

Distance to→ Distance from↓	Country 1	Country 2	Country 3	Country 4	Country 5	Country 6	Country 7	Country 8	Country 9	Country 10
Country 11	10	9	8	7	6	5	4	3	2	1
Country 12	11	10	9	8	7	6	5	4	3	2
Country 13	12	11	10	9	8	7	6	5	4	3
Country 14	13	12	11	10	9	8	7	6	5	4
Country 15	14	13	12	11	10	9	8	7	6	5
Country 16	15	14	13	12	11	10	9	8	7	6
Country 17	16	15	14	13	12	11	10	9	8	7
Country 18	17	16	15	14	13	12	11	10	9	8
Country 19	18	17	16	15	14	13	12	11	10	9
Country 20	19	18	17	16	15	14	13	12	11	10

Notes: See the top half of Table A1 in Appendix A for possible home- and host-country locations that match the distances presented in this table.

Table 2
Distances between the countries in benchmark Sample 2.

Distance to→ Distance from↓	Country A	Country B	Country C	Country D	Country E	Country F	Country G	Country H	Country I	Country J
Country A	0									
Country B	1	0								
Country C	2	1	0							
Country D	3	2	1	0						
Country E	4	3	2	1	0					
Country F	1E+13	1E+13	1E+13	1E+13	1E+13	0				
Country G	1E+13	1E+13	1E+13	1E+13	1E+13	1	0			
Country H	1E+13	1E+13	1E+13	1E+13	1E+13	2	1	0		
Country I	1E+13	1E+13	1E+13	1E+13	1E+13	3	2	1	0	
Country J	1E+13	1E+13	1E+13	1E+13	1E+13	4	3	2	1	0

Notes: See the bottom half of Table A1 in Appendix A for possible home- and host-country locations that match the distances presented in this table.

3.3. Empirical results

To analyze the two benchmark samples, I first consider the amount of country-specific variation in cross-national distance in these samples. The reason is that location-identification is driven by the presence of country-specific variation in measured home-host country distance. In Sample 1, home and host countries together account for almost all (about 100.00%) of variation in measured distance (Table 3, Column 1). Vice versa, in Sample 2, home and host countries together account for almost no (about 0.00%) variation in measured distance (Table 4, Column 1). In similar fashion, there are noticeable differences in average distance to/from the different home and host countries in Sample 1 (Table 3, Column 4) but not those in Sample 2 (Table 4, Column 4).

Next, I turn to the main evidence involving multinomial models with home or host countries as the dependent variables and measured cross-national distance as the independent variable (Tables 3 and 4, Columns 2 and 5). Results confirm that cross-national distance is a statistically highly significant predictor of home or host countries in Sample 1 (Pseudo $R^2 = 0.1471$; $p = 0.0000$; $N = 100$). Hence, it seems that distance can indeed partially identify some specific countries or locations in this sample. The latter finding is demonstrated most clearly by the estimated marginal effects, particularly for host countries 1 and 10 and home countries 11 and 20 ($p = 0.0010$) (Table 3, Column 5). Similarly, results confirm that in a sample with fewer countries that are clear outliers (Sample 2, Table 4), home-host country distance is not a statistically significant predictor of home or host countries (Pseudo $R^2 = 0.0000$; $p = 1.0000$; $N = 90$).

Overall, results strongly support the theoretical claim that distance indicators can partially identify some specific home and/or host countries in a sample. Most notably, results from the multinomial logistic regression models indicate that cross-national distance indicators can partly capture host and/or home country fixed effects when included as an explanatory variable. Hence, it appears that location-identification is not only a theoretical issue. Rather, location-identification is a real phenomenon that can be highly relevant for real-world empirical distance research.

4. Addressing the location-identification problem: towards a pure-distance approach

4.1. Statistically correcting for confounding country-specific variation in measured distance

With an eye to future research, key open question is of course what can be done to minimize the degree to which distance indicators are capturing country fixed effects (and thus overcome the confounded variables problem that appears inherent to distance research). The main challenge posed by cross-national distance indicators partly capturing country fixed effects is a confounded

Table 3

Home- and host-country specific distance, average distances per country and multinomial estimation results for Sample 1: Benchmark case with comparatively much home- and host-country specific variation in measured distance.

% variation in distance that is home- or host-country specific	Categorical dependent variable and multinomial model fit	Country	Average distance to/from country	Marginal probability (percentage point change in likelihood for a one standard deviation increase in distance)
100.00% [p = 0.0000]	Host country Pseudo R ² = 0.1471 p = 0.0000	Country 1	14.5	0.1120 [p = 0.0010]
		Country 2	13.5	0.0806 [p = 0.0096]
		Country 3	12.5	0.0549 [p = 0.0348]
		Country 4	11.5	0.0321 [p = 0.1390]
		Country 5	10.5	0.0106 [p = 0.5825]
		Country 6	9.5	-0.0106 [p = 0.5825]
		Country 7	8.5	-0.0321 [p = 0.1390]
		Country 8	7.5	-0.0549 [p = 0.0348]
		Country 9	6.5	-0.0806 [p = 0.0096]
		Country 10	5.5	-0.1120 [p = 0.0010]
	Home country Pseudo R ² = 0.1471 p = .0000	Country 11	5.5	-0.1120 [p = 0.0010]
		Country 12	6.5	-0.0806 [p = 0.0096]
		Country 13	7.5	-0.0549 [p = 0.0348]
		Country 14	8.5	-0.0321 [p = 0.1390]
		Country 15	9.5	-0.0106 [p = 0.5825]
		Country 16	10.5	0.0106 [p = 0.5825]
		Country 17	11.5	0.0321 [p = 0.1390]
		Country 18	12.5	0.0549 [p = 0.0348]
		Country 19	13.5	0.0806 [p = 0.0096]
		Country 20	14.5	0.1120 [p = 0.0010]

Notes: Sample is Sample 1 in Table 1 and comprises 100 home-host country dyads. Complete estimation results are available on request.

Table 4

Home- and host-country specific distance, average distances per country and multinomial estimation results for Sample 2: Benchmark case with comparatively little home- and host-country specific variation in measured distance.

% variation in distance that is home- or host-country specific	Categorical dependent variable and multinomial model fit	Country	Average distance to/from country	Marginal probability (percentage point change in likelihood for a one standard deviation increase in distance)
0.00% [p = 1.0000]	Host country Pseudo R ² = 0.0000 p = 1.0000	Country A	5.56e + 12	0.0000 [p = 1.0000]
		Country B	5.56e + 12	0.0000 [p = 1.0000]
		Country C	5.56e + 12	-0.0000 [p = 1.0000]
		Country D	5.56e + 12	-0.0000 [p = 1.0000]
		Country E	5.56e + 12	-0.0000 [p = 1.0000]
		Country F	5.56e + 12	-0.0000 [p = 1.0000]
		Country G	5.56e + 12	-0.0000 [p = 1.0000]
		Country H	5.56e + 12	-0.0000 [p = 1.0000]
		Country I	5.56e + 12	0.0000 [p = 1.0000]
		Country J	5.56e + 12	0.0000 [p = 1.0000]
	Home country Pseudo R ² = 0.0000 p = 1.0000	Country A	5.56e + 12	0.0000 [p = 1.0000]
		Country B	5.56e + 12	0.0000 [p = 1.0000]
		Country C	5.56e + 12	-0.0000 [p = 1.0000]
		Country D	5.56e + 12	-0.0000 [p = 1.0000]
		Country E	5.56e + 12	-0.0000 [p = 1.0000]
		Country F	5.56e + 12	-0.0000 [p = 1.0000]
		Country G	5.56e + 12	-0.0000 [p = 1.0000]
		Country H	5.56e + 12	-0.0000 [p = 1.0000]
		Country I	5.56e + 12	0.0000 [p = 1.0000]
		Country J	5.56e + 12	0.0000 [p = 1.0000]

Notes: Sample is Sample 2 in Table 2 and comprises 90 home-host country dyads. Complete estimation results are available on request.

variables problem: one can never be sure what effect or influence is being captured by a distance indicator that has a significant (causal) relationship with the dependent variable of interest (cf. Brouters et al., 2016; Kostova et al., 2019). This problem, in turn, is akin to an omitted variables problem. Hence, it can be addressed, in principle at least, by estimating an empirical model that includes additional variables that control for (unobserved) confounders. Indeed, a standard suggestion for addressing the possible problem of distance-profile conflation is to include home- and host-country fixed effects when estimating a model with cross-national distance as the chief independent variable (e.g., Harzing and Pudelko, 2016). This fixed-effects approach works because country fixed effects absorb all the country-specific variation in the dependent variable that would otherwise be attributed to measured distance. However, controlling for country fixed effects does not directly address the issue that variance in measured distance can contain relatively large amounts of invalid, country-specific variance. In addition, controlling for country fixed effects comes at the expense of being

able to consider possible influences of many country-specific features such as restrictions on inward and outward investments or political risks that are also of substantive interest to researchers and practitioners in the field.

My suggestion therefore is to use an alternative, two-step approach that offers all the advantages of controlling for home- and host-country fixed effects but none of its disadvantages. The first step in this approach is to create an indicator of cross-national distance that is cleansed from confounding home- and host-country influences. Such a pure or corrected distance indicator contains only valid, non-country-specific variance. This first step involves only aggregate-level data on cross-national distance and the home and host countries in one's sample and is empirically straightforward. The only thing that needs to be done to construct a pure indicator of cross-national distance is to estimate a simple OLS model with measured distance as the continuous dependent variable and home- and host-country fixed effects as the independent variables and save the residuals. These residuals provide a corrected measure of continuous dyadic distance that can be matched to any (firm-level) data and used instead of a traditional, uncorrected distance indicator. Hence, in the second step researchers can assess possible effects of cross-national distance in the same way that they would when considering an uncorrected distance indicator. Similar to estimating models that control for home- and host-country fixed effects, the pure-distance approach works for all samples that include at least two home and two host countries. However, a critical advantage of the pure-distance approach is that it is more flexible. The reason is that using distance indicators that are based on residuals retains the possibility of assessing substantive influences of home- and/or host-country characteristics on the phenomenon of interest.

4.2. An empirical illustration of the pure-distance approach and biases due to location-identification

To illustrate the real-world workings of the pure-distance approach, I use it to assess the impact of cross-national distance on a standard aggregate-level dependent variable. This variable is the stock of inward FDI, meaning how much of the FDI stock from a given country is in a particular host country. Data come from the OECD's International Direct Investment Statistics (OECD, 2020) and are for the year 2018. The specific distance metric that I consider is the same indicator of regulative institutional distance as considered by Van Hoorn and Maseland (2016). The underlying country data come from the World Bank's Worldwide Governance Indicators (WGI) project and are also for the year 2018 (World Bank, 2019). Following Van Hoorn and Maseland (2016), I calculate home-host country regulative institutional distance using the principal component of the WGI's four measures of regulative institutions: (i) Rule of Law, (ii) Control of Corruption, (iii) Government Effectiveness, and (iv) Regulatory Quality. The final sample includes 241 home-host country dyads and passes the criterion of minimum seven home and host countries (Van Hoorn and Maseland, 2016). Table A2 in Appendix A presents descriptive statistics and the correlations between the variables that I consider. Similarly, Table A3 in Appendix A shows the home and host countries considered and the extent of location-identification in this sample. Table 5 presents the results.

Model 1 shows the effect of uncorrected regulative institutional distance on inward FDI. From this model, it appears that regulative institutional distance explains about 3.05% of total variation in inward FDI. Models that take into account country-specific variation in measured distance render rather different results, however. Specifically, both the pure-distance approach (Model 2) and the fixed-effects approach (Model 4) indicate a much smaller impact of regulative institutional distance on inward FDI. By

Table 5

An illustrative application of the pure-distance approach and of biases due to location-identification.

Dependent = Inward FDI stock	Model 1	Model 2	Model 3	Model 4	Added country-level explanatory variables		
					Model 5	Model 6	Model 7
Uncorrected regulative institutional distance	-0.1747 [p = 0.0065]	-	-	-0.0546 [p = 0.4705]	-0.0780 [0.2952]	-	-
Pure, corrected regulative institutional distance	-	-0.0546 [p = 0.5238]	-	-	-	-0.0546 [p = 0.5093]	-
Home- and host-country fixed effects included	No	No	Yes	Yes	No	No	No
Home-country regulative institutional profile	-	-	-	-	-0.0145 [p = 0.8292]	0.0122 [p = 0.8452]	0.0122 [p = 0.8450]
Host-country regulative institutional profile	-	-	-	-	0.2397 [p = 0.0007]	0.2730 [p = 0.0000]	0.2730 [p = 0.0000]
No. of observations (= number of dyads)	241	241	241	241	241	241	241
Degrees of freedom regression model	1	1	37	38	3	3	2
Degrees of freedom residuals	239	239	203	202	237	237	238
R ²	3.054%	0.170%	34.04%	34.21%	7.889%	7.631%	7.461%
ΔR ² due to institutional distance	3.054%	0.170%	-	0.170%	-	0.170%	-
ΔR ² due to home- and host-country regulative institutional profile	-	-	-	-	4.035%	7.461%	7.461%

Notes: Table reports estimated coefficients (and p-values) for an OLS regression. As mentioned in the main text, the corrected distance indicator refers to the residuals of a simple OLS regression with measured distance as the continuous dependent variable and home- and host-country fixed effects as the independent variables. The coefficient for distance in Model 1 is a standardized coefficient. This coefficient (and the estimated R² as well) imply a correlation between distance and the dependent variable of -0.1747 (see Table A2 in Appendix A). Table 6 considers the same sample and data and complete data and estimation results are available on request.

mathematical necessity, the estimated coefficient for regulative institutional distance is the same when considering a pure distance measure (Model 2) as when considering an uncorrected distance measure but adding home- and host-country fixed effects as control variables (Model 4). More important, however, is the substantive result that addressing the location-identification problem reduces the explanatory power of regulative institutional distance to about 0.17%. Hence, in this example, location-identification introduces a sizeable upward bias in the estimated impact of cross-national distance.

Meanwhile, the estimated coefficient for regulative institutional distance is not the only coefficient that can be biased because of location-identification. Cross-national distance is a quintessential explanatory variable in international management. Nevertheless, there are also many studies that focus on the effect of various country-level factors on MNEs and that use distance only as a control variable. Moreover, estimated effects for country-level factors are also subject to bias when measured distance partly captures country fixed effects. Models 5–7 therefore consider home- and host-country regulative institutional profile as explanatory variables. Results indicate that considering countries' regulative institutional profile in addition to regulative institutional distance lowers the estimated coefficient for distance quite a bit, also rendering it statistically insignificant at usual levels ($p = 0.2952$ in Model 5 vs. $p = 0.0065$ in Model 1). More importantly, when the location-identification problem is left unaddressed, results could be taken to imply that home- and host-country regulative institutional profile explain some 4% of variation in inward FDI (Model 5 vs. Model 1). Using the pure-distance approach to address the location-identification problem, in contrast, indicates that home- and host-country regulative institutional profile explain some 7.4% of inward FDI (Model 6 and Model 7). Estimated coefficients for home- and host-country regulative institutional profile also change depending on whether or not the problem of location identification is addressed or not. These changes are relatively small however (Model 5 vs. Model 6 and also Model 5 vs. Model 7).²

4.3. Possible limitations of the pure-distance approach

Two features of the pure-distance approach may seem important limitations. The first is the need for a sample that comprises minimum two home and two host countries. The reason is that such samples have been relatively scarce in international management and that collecting additional data to increase the number of home and host countries can be prohibitively expensive. However, as the present paper shows, cross-national distance studies involving samples with fewer than two home or fewer than two host countries are unable to render valid empirical results anyway. The reason is that in such samples all variation in measured distance is (home or host) country-specific and therefore invalid and either all home or all host countries are perfectly identified.

A second, intuitively appealing limitation is that removing all country-specific variation from measured cross-national distance is too blunt and steers estimated effects of cross-national distance towards non-significance. Concerning the first part of this limitation, I should note, however, that using a pure distance indicator is not blunter than adding country fixed effects as control variables. The reason is that both approaches involve absorbing all country-specific variation—the only difference is that, technically speaking, the fixed-effects approach removes all country-specific variation from the dependent variable whereas the pure-distance approach removes all country-specific variation from an independent variable. Concerning the second part of this limitation, I should note that there is, in fact, no theoretical reason why removing country-specific variation in measured distance would imply steering estimated distance effects towards non-significance. Instead, the opposite is equally likely. The underlying mechanism is that the “genuine” distance effect may go in one direction (say a positive effect) whereas the country-specific effects captured by the distance metric may go in the opposing direction (say a negative effect), causing the two effects to cancel out. Hence, there will also be many studies for which an uncorrected distance indicator is insignificant whereas a pure distance indicator would be significant.³

5. Discussion

5.1. The relevance of location-identification *Vis-à-Vis* distance-profile correlations

As with all studies, the analysis presented in this paper may be criticized on various grounds. The most prominent criticism is that the location-identification problem is neither novel nor important. Specifically, the claim would be that location-identification does not add anything beyond potential distance-profile correlations identified in various insightful recent contributions to the distance literature (e.g., [Brouthers et al., 2016](#); [Harzing and Pudelko, 2016](#)). This claim, however, does not do justice to the many ways in which potential location-identification is much more relevant than potential distance-profile conflation is.

Potential distance-profile conflation can occur when distance metrics are effectively a simple difference score between countries on one or more dimensions of countries' environment for doing business.⁴ For instance, institutional distance is usually measured as the absolute difference in the quality of countries' institutions. The U.S. would then tend to rate high in terms of the quality of its institutions and Nigeria would tend to rate much lower. Hence, the institutional distance between the U.S. and Nigeria is high. Open

² Of course, the exact size of the upward or downward biases or the degree over- or understatement of the explanatory power of specific variables due to location-identification is contingent on a variety of factors. Most importantly, biases in estimated effects depend on the exact sample studied, the specific dependent variable considered and the distance indicator used.

³ Following this observation, the location-identification problem can be part of the reasons why the contextual distance literature has often reported small or insignificant effect sizes (see, e.g., the review by [Kirkman et al., 2006](#)).

⁴ I am extremely grateful to an anonymous referee for providing the specific descriptions of distance-profile conflation and the location-identification problem presented in this paragraph and the next. See, also, [Beugelsdijk, Ambos, and Nell \(2018\)](#).

question subsequently is whether the small flow of FDI from the U.S. to Nigeria is due to the high institutional distance between the two countries or due to Nigeria's weak institutions? Both these explanations are plausible. However, because measured institutional distance between the two countries will tend to be highly correlated with these countries' level of institutional quality, it is very hard to distinguish between these two explanations empirically.

The location-identification problem can occur when there is country-specific variation in measured distance. When a country is an outlier on a distance metric, it starts to become uniquely identified (at least within that data set) by the measured distance associated with it. As an unintended consequence, in empirical analyses, the distance metric begins to reflect the impact of other characteristics of one or more countries in the sample in addition to the “true” impact of the distance metric. For example, we might be investigating the impact of institutional distance on the propensity for outward FDI investment. In terms of institutional distance, China could be such an outlier in the sample that the distance to China from all the other countries together uniquely identifies China and begins to capture characteristics of China other than its institutional distance such as the fact that it is a huge market. Of course, few researchers would not include variables such as GDP to control for the size of a country's economy or other obvious and observable confounders. Many other confounding country-level influences, however, are more difficult to identify and control for, if at all. Measurement error in particular is nearly always both partly country-specific and its sources partly unobservable.

Overall, there are three main arguments why the location-identification problem is much more pervasive and fundamental than distance-profile conflation is. First, potential location-identification is a broader phenomenon than possible distance-profile correlations are. Specifically, potential distance-profile correlations are only an issue for a subset of cross-national distance indicators, namely indicators involving differences between countries or so-called contextual distance indicators (Beugelsdijk et al., 2018a).

Second, distance-profile correlations will only lead to a confounded variables problem (i.e., to distance-profile conflation) under certain strict conditions. The ability to identify locations and partly capture country fixed effects, in contrast, is nearly always relevant. To wit, distance-profile conflation will only occur if the following two conditions are met:

- i. Measured distance does in fact correlate significantly with an underlying country profile indicator (e.g., $|\rho| > 0.538$; Van Hoorn and Maseland, 2016, p. 377);
- ii. The underlying country profile dimension has an inherent quality trait associated with it that affects the phenomenon (i.e., the dependent variable) of interest.

Standard distance indicators such as those by Kogut and Singh (1988) and Berry et al. (2010), however, measure home-host country distance on multiple profile dimensions simultaneously and it is not obvious that such combinations of dimensions have a quality trait associated with them that affects many phenomena in international management. Moreover, specific profile dimensions can be expected to affect some phenomena involving MNEs. However, there are also many dependent variables that are not affected by a given country profile dimension. The ability to capture country fixed effects, on the other hand, poses a rather greater risk of a confounded variables problem. The reason is that country fixed effects will not only capture the influence of any possible feature of countries' profiles but also the influence of (unobservable sources of) measurement error in the dependent variable that is specific to a country in one's sample. Hence, country fixed effects are almost certain to pick up a significant country-specific influence, regardless of the phenomenon studied. Fig. 2 summarizes these differences in the occurrence of distance-profile conflation vis-à-vis the location-identification problem. Of course, I thereby limit attention to the subset of distance metrics that are theoretically able to render significant distance-profile correlations in the first place.

Finally, and related to the previous argument, addressing a possible location-identification problem is much more tricky than addressing possible distance-profile conflation is. In case of a distance-profile correlation we know exactly which country-level variable(s) need(s) to be taken into account to avoid the confounded variables problem (i.e., to avoid distance-profile conflation). Specifically, we need to control for the same country variables used to construct the measure of home-host country distance. In case of location-identification, in contrast, we know that there can be a confounded variables problem but not its exact source. Country fixed effects correlate with countless features of countries, including various factors such as country-specific measurement error that

	Strong correlation between home- or host-country profile and phenomenon of interest / the dependent variable	Trivial correlation between home- or host-country profile and phenomenon of interest / the dependent variable
Strong correlation between measured distance and home- or host-country profiles	Distance-profile conflation & location-identification problem	Location-identification problem
Trivial correlation between measured distance and home- or host-country profiles	Location-identification problem	Location-identification problem

Fig. 2. Scenarios for the potential occurrence of a confounded variables problem due to distance-profile correlations or to location-identification for the subset of distance indicators involving differences between countries

Notes: The four scenarios pertain to the subset of cross-national distance indicators that involve differences between home and host countries, e.g., in their environment for doing business. Following Van Hoorn and Maseland (2016, p. 377), strong correlations can be thought of as absolute correlations greater than 0.538 ($|\rho| > 0.538$) and trivial correlations as absolute correlations smaller than 0.176 ($|\rho| < 0.176$). Note, though, that the exact criterion for strong or trivial correlations does not alter the finding that location-identification, as identified in this paper, poses a much greater risk of a confounded variables problem in empirical distance research than possible distance-profile correlations do.

Table 6

An empirical illustration of biases that can remain when a cross-national distance study only addresses the distance-profile conflation problem and not the location-identification problem.

Dependent = Inward FDI stock	Model 8 (=Model 1)	Model 9 (=Model 7)	Model 10	Model 11	Model 12
Uncorrected regulative institutional distance	-0.1747 [<i>p</i> = 0.0065]	-	-	-0.0780 [0.2952]	-
Regulative institutional distance corrected for home- and host-country regulative institutional profile	-	-	-0.0780 [<i>p</i> = 0.3119]	-	-0.0780 [<i>p</i> = 0.2952]
Home-country regulative institutional profile	-	0.0122 [<i>p</i> = .8450]	-	-0.0145 [<i>p</i> = .8292]	0.0122 [<i>p</i> = 0.8450]
Home-country regulative institutional profile	-	0.2730 [<i>p</i> = 0.0000]	-	0.2397 [<i>p</i> = 0.0007]	0.2730 [<i>p</i> = 0.0000]
No. of observations (=number of dyads)	241	241	241	241	241
Degrees of freedom regression model	1	2	1	3	3
Degrees of freedom residuals	239	238	239	237	237
R ²	3.054%	7.461%	0.428%	7.889%	7.889%
ΔR ² due to institutional distance	3.054%	-	0.428%	0.428%	0.428%

Notes: Table reports estimated coefficients (and *p*-values) for an OLS regression, using the same raw data and sample as considered for Table 3. As indicated in the main text, the corrected distance indicator in this table is different from the corrected distance indicator in Table 3. It refers to the residuals of a simple OLS regression with measured distance as the continuous dependent variable and home- and host-country regulative institutional profile scores as the independent variables. By mathematical necessity, the estimated coefficients for the effect of and the percentage of total variance explained by regulative institutional distance are the same for Models 10–12. Complete data and estimation results are available on request.

are difficult to observe, if at all. Moreover, each of these country features may also correlate with the dependent variable and thus give rise to an omitted variable bias (see, also, Fig. 2). Hence, whereas distance-profile conflation is relatively easily avoided—data on the necessary country controls are available by design—the problem of location-identification cannot be addressed by theoretically pinpointing specific country-level sources of omitted variable bias. Table 6 uses the same data as considered for Table 5 to illustrate the (limited) degree to which controlling for relevant home- and host-country profile characteristics helps reduce biases in the estimated impact of cross-national distance. Model 8, which is equal to Model 1 in Table 5, again suggests that regulative institutional distance accounts for some 3.05% of total variation in inward FDI. Taking into account that this estimated distance effect partly reflects direct effects due to home and host countries' regulative institutions, variance explained by regulative institutional distance drops to some 0.428% (Model 10). As before, these results are the same regardless of how exactly these country controls are included, directly or indirectly (i.e., by using the measures of home and host countries' regulative institutional profiles to “correct” the original measure of regulative institutional distance) (Model 11). More important, however, is the result that variance explained continues to be overstated compared to results obtained when all country-level sources of biases are controlled for, 0.428% in Table 6 vs. 0.170% in Table 5 (see, particularly, Model 2 in Table 5). Hence, the empirical evidence confirms the idea that controlling for specific features of countries' institutional profiles reduces the confounded variables problem in empirical distance research. However, as also expected, it does not get rid of all (potential) upward or downward biases and is therefore insufficient for ensuring the validity of an empirical distance study.⁵

5.2. Concluding remarks

This paper has explained and demonstrated how the use of cross-national distance indicators as explanatory variables is bound to suffer a confounded variables problem. Formally, the argument is that measured cross-national distance is typically partly specific to the countries included in one's sample. Country-specific variation in measured home-host country distance can exist irrespective of the number of home/host countries considered (i.e., in large samples) and means that distance indicators partially identify the host and/or home countries in one's sample. Empirically, evidence from multinomial logistic regression models with home or host countries as the categorical dependent variables confirmed cross-national distance indicators' ability to identify some specific host and home locations. Implication is that these indicators partly capture country fixed effects when considered as independent or explanatory variables. Fortunately, biases due to location-identification are rigorously and straightforwardly addressed by using pure distance indicators that are cleansed from confounding home- and host-location influences.

Acknowledgements

The author thanks Vas Taras for his editorial guidance and three anonymous referees for their helpful suggestions. Helpful comments on an earlier draft by Alex Lehr and Roderick Sluiter are gratefully acknowledged.

⁵ As before (see Note 3), the exact size of upward or downward biases that remain when distance-profile correlations are taken into account while location-identification is not, depends on a variety of factors.

Appendix A

Table A1

Possible home and host country locations underlying the home-host distances for benchmark Samples 1 and 2.

Sample	Host countries in the sample		Home countries in the sample	
	Country name	Possible country location	Country name	Possible country location
Sample 1: Benchmark sample created to comprise a comparatively large amount of home- and host-country specific variation in measured distance	Country 1	11	Country	1
	Country 2	12	11	2
	Country 3	13	Country	3
	Country 4	14	12	4
	Country 5	15	Country	5
	Country 6	16	13	6
	Country 7	17	Country	7
	Country 8	18	14	8
	Country 9	19	Country	9
	Country	20	15	10
Sample 2: Benchmark sample created to comprise a comparatively small amount of home- and host-country specific variation in measured distance	Country	1	Country	1
	A	2	A	2
	Country B	3	Country B	3
	Country	4	Country	4
	C	5	C	5
	Country	10,000,000,000,000	Country	10,000,000,000,000
	D	10,000,000,000,001	D	10,000,000,000,001
	Country E	10,000,000,000,002	Country E	10,000,000,000,002
	Country F	10,000,000,000,003	Country F	10,000,000,000,003
	Country	10,000,000,000,004	Country	10,000,000,000,004
	G		G	
	Country		Country	
	H		H	
	Country I		Country I	
Country J		Country J		

Notes: Country locations are purely hypothetical, chosen to render the home-host country distances presented in Tables 1 and 2.

Table A2

Correlations and descriptive statistics for sample and variables considered for the empirical illustration of the pure-distance approach (Tables 5 and 6).

	Correlations				
	Mean and SD	Inward FDI stock (% of total)	Regulative institutional distance	Pure regulative institutional distance	Home-country regulative institutional profile
Inward FDI stock (% of total)	5.26 (0.455)	1			
Regulative institutional distance	0.736 (0.037)	-0.1747	1		
Pure regulative institutional distance	0.0000 (0.028)	-0.0413	0.7560	1	

(continued on next page)

Table A2 (continued)

	Correlations				
	Mean and SD	Inward FDI stock (% of total)	Regulative institutional distance	Pure regulative institutional distance	Home-country regulative institutional profile
Home-country regulative institutional profile	1.26 (0.043)	0.0099	-0.3392	0	1
Host-country regulative institutional profile	1.26 (0.042)	0.2729	-0.4241	0	-0.0084

Notes: Inward FDI stock refers to the percentage of a particular country's total FDI stock that is in a particular host country, e.g., the percentage of Austria's total FDI stock that is in Germany. FDI data are from the [OECD \(2020\)](#) and the sample comprises 241 observations (i.e., home-host country dyads). The main text describes the construction of the measure of regulative institutional distance. The measures of home- and host-country regulative institutional profiles are the principal component of the WGI's four measures of regulative institutions (Rule of Law, Control of Corruption, Government Effectiveness & Regulatory Quality). As indicated in [Section 4.1](#), the pure distance indicator refers to the residuals of an OLS regression with measured regulative institutional distance as the dependent variable and home- and host-country fixed effects as the independent variables. Host countries in the sample are Canada, China, Germany, France, U.K., Italy, Japan and U.S. Home countries in the sample are Australia, Austria, Belgium, Chile, Czech Republic, Denmark, Finland, Greece, Iceland, Ireland, Israel, Italy, Japan, South Korea, Latvia, Lithuania, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, U.K. and U.S. Complete data are available on request.

Table A3

The partial identification of locations by institutional distance in the sample for inward FDI.

% home- or host-country specific variation in regulative institutional distance	Categorical dependent variable and multinomial model fit	Country	Marginal probability (percentage point change in likelihood for a one SD increase in distance)
42.85% [$p = 0.0000$]	Host country Pseudo $R^2 = 0.0479$ $p = 0.0000$	Canada	-0.0062 [$p = 0.7546$]
		China	0.0998 [$p = 0.0000$]
		Germany	-0.0143 [$p = 0.5059$]
		U.K.	-0.0255 [$p = 0.2652$]
		France	-0.0370 [$p = 0.1263$]
		Italy	0.0478 [$p = 0.0050$]
		Japan	-0.0274 [$p = 0.2280$]
		U.S.	-0.0372 [$p = 0.1275$]
		Australia	-0.0096 [$p = 0.4735$]
		Austria	-0.0168 [$p = 0.2942$]
		Belgium	-0.0149 [$p = 0.3301$]
		Chile	-0.0067 [$p = 0.5892$]
		Czech Republic	0.0002 [$p = 0.9881$]
		Denmark	0.0012 [$p = 0.8959$]
		Finland	0.0059 [$p = 0.5265$]
	Greece	0.0183 [$p = 0.0554$]	
	Iceland	-0.0170 [$p = 0.2910$]	
	Ireland	-0.0149 [$p = 0.3247$]	
	Israel	-0.0038 [$p = 0.7245$]	
	Italy	0.0126 [$p = 0.1580$]	
	Japan	-0.0119 [$p = 0.3914$]	
	South Korea	-0.0028 [$p = 0.7987$]	
	Latvia	0.0017 [$p = 0.8663$]	
	Lithuania	0.0008 [$p = 0.9347$]	
	Luxembourg	-0.0010 [$p = 0.9185$]	
	Mexico	0.0410 [$p = 0.0051$]	
	Netherlands	0.0047 [$p = 0.5865$]	
	New Zealand	0.0016 [$p = 0.8723$]	
	Norway	0.0017 [$p = 0.8645$]	
	Poland	0.0073 [$p = 0.4215$]	
Portugal	-0.0026 [$p = 0.8145$]		
Slovakia	0.0084 [$p = 0.3513$]		
Slovenia	0.0000 [$p = 0.9973$]		
Spain	0.0002 [$p = 0.9872$]		
Sweden	0.0012 [$p = 0.9058$]		
Switzerland	0.0025 [$p = 0.7962$]		
Turkey	0.0296 [$p = 0.0146$]		
U.K.	-0.0162 [$p = 0.3054$]		
U.S.	-0.0207 [$p = 0.2387$]		
	Home country Pseudo $R^2 = 0.0313$ $p = .0000$		

Notes: Sample comprises 241 home-host country dyads (see [Table A2](#)). Complete data and estimation results are available on request.

References

- Ambos, B., Håkanson, L., 2014. The concept of distance in international management research. *J. Int. Manag.* 20 (1), 1–7.
- Bae, J.-H., Salomon, R., 2010. Institutional distance in international business research. *Advances in International Management* 23, 327–349.
- Berry, H., Guillén, M.F., Zhou, N., 2010. An institutional approach to cross-national distance. *J. Int. Bus. Stud.* 41 (9), 1–21.
- Beugelsdijk, S., Maseland, R., Onrust, M., Van Hoorn, A., Slangen, A., 2015. Cultural distance in international business and management: from mean-based to variance-based measures. *Int. J. Hum. Resour. Manag.* 26 (2), 165–191.
- Beugelsdijk, S., Kostova, T., Roth, K., 2017. An overview of Hofstede-inspired country-level culture research in international business since 2006. *J. Int. Bus. Stud.* 48 (1), 30–47.
- Beugelsdijk, S., Ambos, B., Nell, P.C., 2018a. Conceptualizing and measuring distance in international business research: recurring questions and best practice guidelines. *J. Int. Bus. Stud.* 48 (9), 1113–1137.
- Beugelsdijk, S., Kostova, T., Kunst, V.E., Spadafora, E., van Essen, M., 2018b. Cultural distance and firm internationalization: a meta-analytical review and theoretical implications. *J. Manag.* 44 (1), 89–130.
- Brouthers, L.E., Marshall, V.B., Keig, D.L., 2016. Solving the single country sample problem in cultural distance studies. *J. Int. Bus. Stud.* 47 (4), 471–479.
- Chikhouni, A., Edwards, G., Farashahi, M., 2017. Psychic distance and ownership in acquisitions: direction matters. *J. Int. Manag.* 23 (1), 32–42.
- Estrin, S., Nielsen, B.B., Nielsen, S., 2017. Emerging market multinational companies and internationalization: the role of home country urbanization. *J. Int. Manag.* 23 (3), 326–339.
- Geleilate, J.M.G., Magnusson, P., Parente, R.C., Alvarado-Vargas, M.J., 2016. Home country institutional effects on the multinationality-performance relationship: a comparison between emerging and developed market multinationals. *J. Int. Manag.* 22 (4), 380–402.
- Ghemawat, P., 2016. *The Laws of Globalization and Business Applications*. Cambridge University Press, Cambridge.
- Harzing, A.-W., 2003. The role of culture in entry-mode studies: from neglect to myopia? *Advances in International Management* 15, 75–127.
- Harzing, A.-W., Pudelko, M., 2016. Do we need to distance ourselves from the distance concept? Why home and host country context might matter more than (cultural) distance. *Manag. Int. Rev.* 56 (1), 1–34.
- Huang, Z., Zhu, H.S., Brass, D.J., 2017. Cross-border acquisitions and the asymmetric effect of power distance value difference on long-term post-acquisition performance. *Strateg. Manag. J.* 38 (4), 972–991.
- Hutzschenreuter, T., Kleindienst, I., Lange, S., 2016. The concept of distance in international business research: a review and research agenda. *Int. J. Manag. Rev.* 18 (2), 160–179.
- Kirkman, B.L., Lowe, K.B., Gibson, C.B., 2006. A quarter century of culture's consequences: a review of empirical research incorporating Hofstede's cultural values framework. *J. Int. Bus. Stud.* 37 (3), 285–320.
- Kogut, B., Singh, H., 1988. The effect of national culture on the choice of entry mode. *J. Int. Bus. Stud.* 19 (3), 411–432.
- Konara, P., Shirodkar, V., 2018. Regulatory institutional distance and MNCs' subsidiary performance: climbing up vs. climbing down the institutional ladder. *J. Int. Manag.* 24 (4), 333–347.
- Kostova, T., Beugelsdijk, S., Scott, W.R., Kunst, V.E., Chua, C.H., van Essen, M., 2019. The construct of institutional distance through the lens of different institutional perspectives: review, analysis, and recommendations. *J. Int. Bus. Stud.* 51, 467–497. <https://doi.org/10.1057/s41267-019-00294-w>.
- Long, S.J., Freese, J., 2014. *Regression Models for Categorical Dependent Variables Using Stata*, 3rd edn. Stata Press, College Station.
- López-Duarte, C., Vidal-Suárez, M.M., González-Díaz, B., 2019. Cross-national distance and international business: an analysis of the most influential recent models. *Scientometrics* 121 (1), 173–208.
- OECD, 2020. *Inward FDI Stocks by Partner Country (Indicator)*. <https://doi.org/10.1787/a1818a82-en>. (Accessed on 08 April 2020).
- Shenkar, O., 2001. Cultural distance revisited towards a more rigorous conceptualization and measurement of cultural differences. *J. Int. Bus. Stud.* 32 (3), 519–535.
- Tung, R.L., Verbeke, A., 2010. Beyond Hofstede and GLOBE: improving the quality of cross-cultural research. *J. Int. Bus. Stud.* 41 (8), 1259–1274.
- Van Hoorn, A., Maseland, R., 2014. Is distance the same across cultures? A measurement-equivalence perspective on the cultural distance paradox. In: Verbeke, A., Van Tulder, R., Lundan, S. (Eds.), *Multinational Enterprises, Markets and Institutional Diversity*. Howard House: Emerald, pp. 207–227.
- Van Hoorn, A., Maseland, R., 2016. How institutions matter for international business: institutional distance effects vs institutional profile effects. *J. Int. Bus. Stud.* 47 (2), 374–381.
- Verbeke, A., Van Tulder, R., Puck, J., 2017. Distance in international business studies: concept, cost and value. In: Verbeke, A., Puck, J., Van Tulder, R. (Eds.), *Distance in International Business: Concept, Cost and Value*. Howard House: Emerald, pp. 17–43.
- World Bank, 2019. *Worldwide governance indicators project*. www.govindicators.org.
- Yeganeh, H., 2014. A weighted, mahalanobian, and asymmetrical approach to calculating national cultural distance. *J. Int. Manag.* 20 (4), 436–463.
- Zaheer, S., Schomaker, M.S., Nachum, L., 2012. Distance without direction: restoring credibility to a much-loved construct. *J. Int. Bus. Stud.* 43 (1), 18–27.