Journal of World Business xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of World Business

journal homepage: www.elsevier.com/locate/jwb

The selection, use, and reporting of control variables in international business research: A review and recommendations

Bo Bernhard Nielsen[a,b,*], Arpit Raswant[a]

[a] The University of Sydney Business School, Discipline of International Business, The University of Sydney, Australia
[b] Department of Strategic Management and Globalization, Copenhagen Business School, Denmark

ARTICLE INFO

Keywords:
Control variable
Statistical control
Research method
Quantitative method
Review
Recommendation
Selection
Use
Reporting
Validation
International business

ABSTRACT

This study explores the selection, use, and reporting of control variables in studies published in the leading international business (IB) research journals. We review a sample of 246 empirical studies published in the top five IB journals over the period 2012–2015 with particular emphasis on selection, use, and reporting of controls. Approximately 83% of studies included only half of what we consider Minimum Standard of Practice with regards to controls, whereas only 38% of the studies met the 75% threshold. We provide recommendations on how to effectively identify, use and report controls in IB studies.

1. Introduction

Control variables (CVs) constitute a central element of the research design of any empirical study. Confounding variables are likely to covary with the hypothesized focal independent variables thus limiting both the elucidation of causal inference as well as the explanatory power of the model (Stone-Romero, 2009; Pehazur & Schmelkin, 1991). Therefore, researchers must seek to rule out threats to valid inferences in order to determine to what extent the focal independent variables behave as hypothesized. This is typically done by including (controlling for) extraneous variables that are deemed theoretically (or empirically) important but are not focal variables of the study (Kish, 1959). The literature sometimes refers to such variables as covariates, confounding variables, nuisance variables, control variables or simply controls (Atinc, Simmering, & Kroll, 2012; Breaugh, 2008). Researchers need to account for these variables either through experimental design (before the data gathering) or through statistical analysis (after the data gathering process). In this way the researchers are said to account for their effects to avoid a false positive (Type I) error (i.e. falsely concluding that the dependent variables are in a causal relationship with the independent variable). Inadequate attention to controls is a major threat to the validity of inferences made about cause and effect (internal validity).[1]

One way of controlling by inclusion is to use a matched-group design where particular entities (e.g., state-owned and privately owned firms) that vary in terms of independent and dependent variables are matched on specific criteria (Estrin, Meyer, Nielsen, & Nielsen, 2016). An alternative way of controlling is exclusion by holding particular variables constant, such as limiting a study to emerging market firms only (Buckley, Elia, & Kafouros, 2014). Yet the most common way to control for extraneous influences is via statistical controls. Statistical controls aim at identifying potential sources of influence during study design and including CVs representing these sources of influence during data collection. During data analysis, researchers then control for these extraneous effects by mathematically partialling out variance associated with CVs in calculating relationships between other variables, thereby reducing the risk of Type II errors (Carlson & Wu, 2012; Spector, Zapf, Chen, & Frese, 2000). In this study we focus on IB research that includes statistical controls as non-hypothesized variables in regression type studies.

When regressing for instance firm performance (or entry mode) on other variables, IB researchers attempt to establish which specific variables influence the prediction and which do not. This is typically done by considering whether each variable's contribution remains

* Corresponding author.
E-mail addresses: bo.nielsen@sydney.edu.au (B.B. Nielsen), arpit.raswant@sydney.edu.au (A. Raswant).

[1] We note that causal inferences in IB studies are difficult to make and point to the excellent Editorial in JIBS by Reeb and colleagues (Reeb, Sakakibara, & Mahmood, 2012) on endogeneity in international business research for recommendation as to how to best approximate randomized-controlled experiments.

https://doi.org/10.1016/j.jwb.2018.05.003
Received 10 November 2017; Received in revised form 14 April 2018; Accepted 14 May 2018

statistically significant after controlling for other predictors. In multiple regression, when the coefficient of a predictor variable differs significantly from zero, most scholars conclude that this variable makes a "unique" contribution to the outcome. CVs are assumed to be confounding, that is, producing distortions in observed relationships. For this reason, researchers typically clearly assign some variables as being merely controls, or variables of no particular theoretical interest, that need to be somehow removed in their effects on the study. While statistical controls *are* able to adjust relationships between variables for the action of other variables, this ability is based on certain implicit assumptions about the underlying role of control variables on either the observed measures or the underlying constructs of interest. More generally, the argument seems to be that we decrease the aggregate bias for every additional relevant variable that we include. The inefficiency part of the equation is, however, rarely mentioned, as control variables often do have real effects. Yet, the mathematics of regression analysis do not support the argument that more variables in a regression, even relevant ones, necessarily makes the regression results more accurate (Clarke, 2005). In fact, even small amounts of measurement error in control variables "are magnified as more variables are added to the equation in an attempt to control for other possible sources of bias." (Griliches, 1977: 12).

Control variables are of extreme importance in econometric analyses for a number of reasons. First, the variables included in the analysis drive the results of any statistical analysis of data. Hence, the improper use (inclusion or exclusion) of CVs may distort results and produce misleading findings. Similar to any other variable included in a model (e.g., any predictor or criteria variable), decisions regarding which controls to include affect the significance levels and estimated effect sizes of the other variables. Second, replication and generalizability of results cannot be done without specific knowledge of which factors were controlled, the measurement of these controls, and the specific method utilized for controlling. Finally, inadequate justification and reporting of controls render any extension difficult. This includes meta-analyses, which cannot be conducted on studies where controls are unknown, unjustified, or measurement and descriptive statistics are not reported. In order to advance IB research and build a cumulative body of knowledge about certain phenomena, the correct selection, inclusion and treatment as well as documentation and reporting of CVs is critical since controls often serve as inspiration for new studies of relationships (i.e., as potential moderators/mediators, IVs or even DVs).

We build upon insights from previous articles on the role of control variables in social science research (e.g., Atinc et al., 2012; Becker, 2005; Becker et al., 2015; Bernerth & Aguinis, 2015; Breaugh, 2006; Carlson & Wu, 2012; Spector & Brannick, 2011). These studies document the (mis)use of control variables in social science research by analyzing how published work in the top tier management and organizational psychology journals have treated controls inadequately. To the best of our knowledge, however, this is the first comprehensive review of the selection, use, and reporting of control variables in IB research (also see Aguinis, Cascio, & Ramani, 2017). As such, we join an important (recent) conversation within the IB research community which calls for more attention to both methodological rigor in empirical testing and preciseness in presentation and reporting of results (e.g., Andersson, Cuervo-Cazurra, & Nielsen, 2014; Ahlstrom, 2015; Cortina, Köhler & Nielsen, 2015; Cuervo-Cazurra, Andersson, Brannen, Nielsen, & Reuber, 2016; Kingsley, Noordewier, & Bergh, 2017; Welch & Piekkari, 2017).

IB research is particularly vulnerable to issues arising from poor treatment in terms of selection, analysis and reporting of control variables due to its complex and multi-disciplinary nature, often spanning multiple countries and contexts (Aguinis et al., 2017; Cuervo-Cazurra et al., 2016). IB studies involve phenomena where country level context (e.g., institutional or cultural) often play a decisive role as boundary conditions for theory development. In fact, what sets IB studies apart from more general strategy, management or organizational research is the cross-border (international) business context in which actors (individuals, teams, firms or even industries) act and interact (Zaheer, Schomaker, & Nachum, 2012). This international context has important implications for use of control variables as it helps establish the boundaries of applicability surrounding a particular empirical argument and rule out alternative or confounding explanations of findings (Teagarden, Von Glinow, & Mellahi, 2018). As noted by Cho and Padmanabhan (2005: 309) "no international business study can be complete unless there is an explicit variable controlling for cultural distance."

This study seeks to investigate the state-of-the-art of treatment of control variables in IB studies. For comparison reasons, we focus on specific issues pertaining to the selection, use and reporting of control variables studied previously, but re-interpret these in terms of specific importance to IB research. Together with our concrete recommendations, this approach is intended to provide IB scholars with a comprehensive yet easy to follow guide to improve their treatment of control variables. In addition, we specifically examine the treatment of country level, contextual variables as controls in IB research, and recommend ways to improve practice with regards to such controls.

We start by introducing our sample and method followed by a thorough analysis of the current CV use and reporting in 246 empirical articles published in the top five IB journals during the period 2012–2015. We compare and contrast the use of controls both between the five IB journals and with result from previous studies in other fields. Based on our findings, we provide a set of recommendation to guide future authors, reviewers and editors toward a more consistent and accurate way of controlling for extraneous variables in IB research.

2. Method

2.1. Selection of articles

In an attempt to be comprehensive, we coded all empirical articles published in five top IB journals over the period 2012–2015 with regards to the use and reporting of controls. Given our focus on use and reporting of statistical extraneous CVs,[2] we did not evaluate non-empirical studies, editorials or research forums, qualitative studies, simulations or experimental studies. To further ensure clarity and comparability, we omitted longitudinal panel data studies and studies using SEM, GMM, multilevel, 3-stage least square, *meta*-analysis or other methods where use of control variables is less equipollent. We also omit studies where the use of control variables could not easily be discerned from information provided, including studies with no correlation table and articles with no control variables at all. In articles that reported multiple studies with different controls (often as robustness checks), we focused on the primary test of hypotheses and thus treated it as a single article. Our final sample consisted of 246 articles published in *Journal of World Business (JWB – 54 articles), Journal of International Business Studies (JIBS – 44 articles), Journal of International Management (JIM – 34 articles), Management International Review (MIR – 35 articles),* and *International Business Review (IBR – 79 articles)*.

2.2. Coding of articles

Building on insights from prior research on controls in empirical research, we developed a coding scheme designed to identify both

[2] Achieving appropriate statistical control depends on the researcher's intent; different objectives such as a) "purification" of relationship between variables, b) estimating a "controlled" relationship between two variables that accounts for the effects of other meaningful variables, or (c) determining the "incremental" contribution that a variable makes to the prediction of a DV after the effects of other variables have been considered may require different types of information and analytical strategies (Carlson & Wu, 2012; Conger & Jackson, 1972). Where appropriate, we discuss these differences in relation to our recommendations.

Table 1
Control Variable Dimensions, Reasoning and Coding.

Criterion	#	CV Dimension	Conceptual Reasoning	Coding
Selection	CD1	Justification for inclusion of control variables	Providing theoretical basis for inclusion of controls and/or prior empirical evidence is key to including relevant controls (Atinc et al., 2012; Becker, 2005; Bernerth & Aguinis, 2015)	Coded 1 if explanation was missing for at least one CV; otherwise 0;
	CD2	Justification for measurement of CVs	The nature of measurement of controls is key to evaluating a study's validity and reliability (Becker, 2005; Breaugh, 2006)	Coded 1 if measurement was missing for at least one CV; otherwise 0
	CD3	Justification for choice of CVs based on citation	Citations in support of *the use of* all CVs help establish historic relevance (Atinc et al., 2012; Bernerth & Aguinis, 2015)	Coded 1 if references were missing for at least one CV; otherwise 0
Use	CD4	Theoretical prediction of directionality of CVs	Prediction of sign (positive or negative) for any CV-dependent variable relationship is key to understanding potential impact of inclusion (Becker, 2005; Carlson & Wu, 2012: Bernerth & Aguinis, 2015)	Coded 1 if prediction of sign was missing between at least one CV and the dependent variable; otherwise 0
	CD5	Use of CVs with no association to dependent variable(s)	Including controls that are not associated with dependent variable(s) may reduce power to detect meaningful relationships (Becker, 2005; Breaugh, 2006, Carlson & Wu, 2012)	Coded 1 if none of the CVs were significantly correlated with DV; otherwise 0
	CD6	Use of CVs with unknown, high correlation with other independent variables	Including controls with unknown, high correlations with other independent variables may confound statistical control efforts, potentially disrupting efforts to purify relationships and further restricting residuals (Carlson & Wu, 2012; Spector & Brannick, 2011)	Coded 1 if any of the CVs were correlated above 0.5 with other independent variables; otherwise 0
Reporting	CD7	Reporting of which independent variables are treated as controls	A separate section dedicated to CVs − either with the header Control Variables or a specific paragraph discussing explicitly the included control variables, their justification, measurement and relevance − indicate more attention to these variables and provide easy access to pertinent information (Atinc et al., 2012)	Coded 1 if the article did not have a specific control variable section; otherwise 0
	CD8	Reporting of descriptive statistics of CVs (A)	Reporting both means and standard deviations of all controls provides key information about the properties of the sample (Becker, 2005)	Coded 1 if mean or/and standard deviation was missing for at least one CV; otherwise 0
	CD9	Reporting of descriptive statistics of CVs (B)	Reporting the range and/or minimum and maximum value of each control variable establishes how widely spread out the most extreme observations are	Coded 1 if range was missing for at least one CV; otherwise 0
	CD10	Reporting of CVs in correlation table	Including ALL controls in the correlation table show associations between controls and other variables (Becker, 2005)	Coded 1 if at least one CV was missing in correlation table; otherwise 0
	CD11	Reporting of results (A): Model with controls only	A base model with only controls allows for incremental prediction as subsequent models provide evidence of utility of other independent variables (Carlson & Wu, 2012)	Coded 1 if regression model with controls only was missing; otherwise 0
	CD12	Reporting of results (B): Model without controls	A model which includes all independent variables but no controls may provide important information about the utility of IVs to explain un-controlled variation in dependent variable(s) (Becker et al., 2015; Breaugh, 2006, 2008)	Coded 1 if regression model without controls was missing; otherwise 0
	CD13	Discussion of controls	Control variables, whether their coefficients are significant or not, provide important information about the reliability, generalizability and replicability of results (Atinc et al., 2012; Becker, 2005)	Coded 1 if there was no explicit discussion of the influence of any controls; otherwise 0
Context	CCV1	Inclusion of contextual CV (CCV) in multi-country study	Multi-country studies need to account for country level contextual variance (Makino, Isobe, & Chan, 2004; Tong, Alessandri, Reuer, & Chintakananda, 2008)	Coded 1 if study does not include contextual CV; otherwise 0
	CCV2	Use of country dummy CCV in multi-country study	Country dummies represent a simple (yet inadequate) CCV (Nielsen & Nielsen, 2010)	Coded 1 if study does not include country dummy CCV; otherwise 0
	CCV3	Use of absolute CCV in multi-country study	Absolute controls, such as GDP, GDP per capita, unemployment rate, institutional quality, etc. in either home or host country represent a CCV (Duanmu, 2014)	Coded 1 if study does not include absolute CCV; otherwise 0
	CCV4	Use of Hofstede-based CCV in multi-country study	Despite ample criticism, Hofstede-based CCVs remains popular in IB studies (Shenkar, 2012; Javidan et al., 2006)	Coded 1 if study does not include Hofstede-based CCV; otherwise 0
	CCV5	Use of relative CCV in multi-country study	Relative (i.e. distance index) controls, such as Kogut and Singh (1988) represent a *relative* CCV (Ertug, Cuypers, Noorderhaven, & Bensaou, 2013)	Coded 1 if study does not include relative CCV; otherwise 0

positive and negative treatment of controls in IB research. Following best-practice-recommendations offered by Duriau, Reger, and Pfarrer (2007), the authors started by carefully reading the selected articles from each of the five journals. Based on these initial readings, we developed a list of coding categories and coded the studies. After several revisions and using a combination of open and axial coding (Miles & Huberman, 1994; Strauss & Corbin, 1998), we selected 13 recognized dimensions upon which we coded each study to assess the overall treatment of controls in IB studies. The dimensions were designed to address issues pertaining to justification for inclusion of CVs, including inclusion of impotent controls that are uncorrelated with the dependent variable, description of measurement of CVs, hypothesized effects of CVs, reporting of descriptive statistics and relationships with other variables in correlation table, reporting and discussion of influence of controls on primary relationships. After careful deliberation and to ensure comparability, we followed previous studies (e.g., Atinc et al., 2012; Bernerth & Aguinis, 2015; Carlson & Wu, 2012) in using dichotomous (0/1) coding according to whether or not a study followed best practices on a given dimension. Dichotomizing a potentially continuous variable may result in an inflation of Type I errors as studies are forced into either of two extremes. However, given the nature of our dimensions and the intention of our study to provide best practice recommendations, we did not consider this a problem. Nevertheless, to appease concerns about bias due to outliers in our results, we proceeded to draw a random sample and recode two dimensions (CD1 and CD2) according to their underlying properties (i.e. the actual number of studies that omitted these). Next, we applied three different thresholds according to 'missingness'; 50%, 33%, and 25% in order to test the robustness of our findings. The results (not reported here) clearly supported our initial coding; for CD1 we found that 85% of the studies which we had coded as missing (1) explained less than 50% of the control variables included in their study. For lower thresholds, the results on CD1 showed that 95% of the studies are missing explanation for 33% of control variables in their study, while 100% of studies are missing explanation of at least 25% of the control variables utilized in their study. For CD2 we see similar results; 55% of studies we coded as missing if description of measurement of controls (1) did so for 50% or more of the control variables. For lower thresholds, 65% of studies were missing measurements for at least 33% of their controls and 75% were missing measurements for at least 25% of controls. We strongly believe that researchers should strive for no missingness on all the dimensions as lowering the threshold constitutes a methodological compromise. Hence, our binary coding is designed to capture the rigor in relation to selection, use, and reporting of control variables which is required to ensure methodological rigor.

In addition to these recognized dimensions of control variables, we also coded a specific new category of variables – which we call "Contextual Control Variables". IB studies are characterized by the cross-country context in which business activities take place. One of the most important and unique dimensions of IB context − country level variation – provides important insights into both the quality, reliability and appropriateness of multi-country IB studies (Teagarden et al., 2018). Hence, we proceeded to code the nature of contextual control variables in the following way: First we coded all studies according to whether or not they were single-country studies (coded 1) or multi-country studies (coded 0). Only the latter were retained (n = 139) as country level context was deemed particularly important in such IB studies. Next, we coded these 139 studies according to whether they included any (n = 107) country level contextual controls at all. This set of studies were then scrutinized further and coded according to the type and nature of control variables included. We coded these country level controls according to a hierarchy of appropriateness of measures, starting with simple country level dummies (coded 0 if included, otherwise 1). Next, we coded studies according to inclusion of *absolute* country level controls, such as GDP, GDP per capita, unemployment rate, institutional quality, etc. in either home or host country (coded 0 if included, otherwise 1). While such absolute measures of home or host country context arguably are better than simple dummies, research suggests that it is really the *relative distance* between home and host country contextual environments that matter in IB (Caprar, Devinney, Kirkman, & Caligiuri, 2015). Given the preponderance of studies using Hofstede (Hofstede, 1980) derived measures of distance and the ongoing debate/criticism of this approach (e.g., Javidan, House, Dorfman, Hanges, & De Luque, 2006; Kirkman, Lowe, & Gibson, 2006), we first coded studies for reliance on Hofstede (either as absolute country scores or as relative distance measures). Next, we coded studies according to inclusion of specific relative country distance types of controls, such as Kogut/Singh measures of cultural distance or measures of institutional or administrative distance, for instance based on the CAGE framework etc. (coded 0 if included, otherwise 1). We also coded to what extent such relative distance measures were uni-directional (i.e. from home to host country only, or vice versa) or multi-directional (i.e. looked at from both directions) (Schaffer & Riordan, 2003). Specifically, we coded the dimensions as shown in Table 1.

We initially coded a subset of articles from each journal separately to assess the level of agreement across coding dimensions. General agreement was around 90% and we discussed the few categories where there were discrepancy and refined our dimensions to ensure higher reliability in coding. We then divided the articles approximately evenly between us and coded the remaining articles. Upon completion of coding, we randomly selected 20 articles, 10 from each coder's sample, and recoded them independently. For these 20 articles our agreement level across each dimension ranged from 90 to 100%, with an average of above 95%. We thus conclude that the coding of the two authors was highly consistent across all dimensions.

Table 2
Percentage of Studies Incorporating and Reporting Controls Inadequately.

Criterion	Dimension/Journal	JWB	JIBS	JIM	MIR	IBR	Total Average	Atinc et al. (2012)
Selection	CD1	66.67	54.55	47.06	57.14	49.37	**54.88**	52.0
	CD2	24.07	25.00	20.59	0.00	8.86	**15.45**	10.0
	CD3	68.52	52.27	50.00	48.57	59.49	**57.32**	80.3
Use	CD4	85.19	81.82	88.24	74.29	91.14	**85.37**	91.2
	CD5	31.48	25.00	29.41	28.57	39.24	**32.11**	N/A
	CD6	18.52	45.45	64.71	20.00	48.10	**39.43**	N/A
Reporting	CD7	33.33	40.91	35.29	37.14	32.91	**35.37**	24.7
	CD8	14.81	11.36	8.82	5.71	10.13	**10.57**	13.7
	CD9	87.04	75.00	70.59	82.86	74.68	**78.05**	N/A
	CD10	12.96	29.55	8.82	17.14	8.86	**14.63**	10.5
	CD11	38.89	43.18	29.41	51.43	40.15	**40.65**	N/A
	CD12	96.30	100.00	97.06	100.00	92.41	**96.34**	N/A
	CD13	51.85	43.18	38.24	28.57	27.85	**37.40**	73.3

3. Findings

Table 2 reports the percentages of each of the dimensions used to assess the incorporation and reporting of CVs in IB studies. The table illustrates the differences in practices with regards to use and presentation of CVs across IB journals and on average with regards to both individual and aggregated control dimensions.

We begin by discussing the overall trends and compare results to findings from four top management journals (*Academy of Management Journal, Strategic Management Journal, Journal of Management, and Journal of Applied Psychology*) as reported by Atinc et al. (2012). The column labeled 'Mean' on the right reports the average percentage of all studies in the five IB journal (across all four years) which are not incorporating and reporting controls appropriately on each of the dimensions outlined in Table 1. Comparing these results to the average across the sample of four top management journals reported by Atinc et al. (2012), we observe rather similar results on several dimensions (e.g., CD1, CD8, and CD10).[3]

On a few dimensions, however, IB journals seem to be doing either better or worse compared to the top four management journals. First, we note with great interest that authors in IB journals appear to do a better job of justifying the use of controls by citations (CD3) to other empirical studies that previously have utilized these (approx. 43% versus only about 20% in the management journals). Looking across the IB journals, this seems in part to be driven by better use of citations in MIR, JIM and JIBS, though all IB journals outperform the management journal average. By the same token, IB journals also appear to be, on average, better at reporting both means and standard deviations of all controls (CD8) compared to the management journals (almost 90% for IB journals compared to just over 86% for management journals). While most IB journals' average is better than the average for the management journals, MIR and JIM seem to be particularly vigilant in reporting sample properties such as means and standard deviations, whereas JWB seemingly is slightly underperforming. Finally, while one would expect far better, we find some comfort in the high proportion of IB studies (well above 60%) that explicitly discuss the influence of controls in their results and/or discussion sections (CD13) compared to the top four management journals (only slightly above 25%). On the other hand, we observe that many IB studies (approx. 35%) do not provide a separate section discussing control variables (CD7), whereas this seems more common practice in the management journals, where only about 25% omit such a section. We also note that, while the means are relatively similar, several prominent IB journals do not appear to provide adequate theoretical justification for inclusion of particular controls.

For several of the dimensions coded in our study, no comparison exists in the study by Atinc et al. (2012). Other studies, however, provide relatively comparable dimensions on some of these dimensions. For instance, Carlson and Wu (2012) analyzed control variable practice in *Academy of Management Journal*, *Journal of Applied Psychology,* and *Strategic Management Journal* for studies published during 2007. They coded studies according to whether they justified their selection (inclusion) of controls based on their relationship with the dependent variable(s) and found 86% to do so. Similarly, we coded our sample studies for the extent to which controls were associated with the dependent variable(s) (CD5) and found that, on average, only less than 68% of studies could justify inclusion based on this relationship. Moreover, similar to Carlson and Wu (2012), we also coded our studies for high (above 0.5) correlations between CVs and other independent variables (CD6). While high (0.5 or even 0.7) correlation between CVs and IVs does not equate to multicollinearity per se, it may confound regression results and thus should act as a "flag" to researchers to conduct further analysis (i.e. VIFs) to rule out such issues.[4] We found more than 39% of IB studies to suffer from this issue. While this compares favorably, on average, to the three journals analyzed by Carlson and Wu (2012), which found just above 40% of studies to exhibit high correlations between CVs and independent variables, we note a high variability between IB journals on this particular dimension. While MIR and IBR appear relatively close to the average, JWB excels with only slightly about 18.5% of studies across the four years exhibiting high correlations. On the other hand, both JIBS (~45%) and in particular JIM (~65%) seem to be plagued by such high correlations. Finally, Carlson and Wu (2012) also analyzed the practice of reporting a base model with controls only in the regression table (CD11). Their findings revealed that almost all studies (96%) in the three journals (in the year 2007) reported an 'empty' model with CVs first in hierarchical analyses. For our sample of IB journals, however, only ~59% of studies used this practice.

We also note that almost no studies (less than 5%) report a final model with all independent variables but no controls (CD12) as suggested by several scholars as good practice (e.g., Becker et al., 2015; Breaugh, 2006, 2008). We also coded for the reporting of range and/or minimum and maximum values in descriptive statistics (CD9) and found less than a quarter (~22%) to do so. Finally, we found that almost 31% of IB studies did not report standard errors of controls (or indeed any variables) in their regression tables (not reported in Table 1), clearly a practice that leaves much to be desired. Standard errors provide important information about the quality of the estimates of the mean. Standard errors also allow for calculations of significance levels. Related to this, we also found that extremely few (less than 2%) of IB studies reported effect sizes when discussing results; a practice now made mandatory for studies published in *Strategic Management Journal* and other top journals.

3.1. Country level controls in IB studies

In addition to these comparative dimensions of controls, we also analyzed a new set of controls of particular relevance and importance to IB studies. These controls (CCV1-CCV5) relate specifically to country level contextual factors, the influence of which is deemed essential to tease out particularly in cross-country studies. Table 3 shows the percentage of studies in our sample that use various approaches to control for contextual environment in multi-country studies.

In our sample of 246 studies, 139 were focusing on more than one country explicitly. Of these studies, about 77% (107 studies) included at least one contextual control variable (CCV1). Looking more closely at these 107 studies, we proceeded to examine the specific type of contextual control used. As indicated in Table 3, more than 37% of the studies that included at least one CCV utilized a simple country dummy as the only means of controlling for contextual variance. Similarly, another 38% of studies used only absolute measures of country factors, such as GDP, GDP per capita or quality of institutions etc. as CCV. Although Hofstede's cultural dimensions continue to be popular among IB studies (27%), less than 14% of the studies employed some kind of relative measure of distance between home and host countries. We also note (not reported in Table 3) that none of the studies make use of multi-directional relative measures capturing the distance between home and host countries from both country's perspective (Schaffer & Riordan, 2003). Looking at the variance between IB journals in the use

[3] Since we are comparing four years (2012–2015) of observations across five top IB journals with five years (2005–2009) of observations across four top management journals, caution must be taken in interpreting the results. We also caution that the number of observations within each journal is relatively small.

[4] Variance Inflation Factor (VIF) analysis is often conducted to rule out multicollinearity issues, however, the focus of such analysis is typically on the variance of a regression coefficient and the overall stability of results. The correlation of controls with other independent variables, in addition to sample size, model complexity and a number of other substantive issues are equally important to rule out multicollinearity (Goldberger, 1991). None of the studies in our sample that reported 0.5 or above correlations between CVs and IVs conducted specific VIFs on CVs.

Table 3
Results: Percentage of studies and their treatment of contextual control variables.

Journal	CCV1 Inclusion of CCV (%)	CCV2 Only Country Dummy	CCV3 Only Absolute CCV	CCV4 Hofstede	CCV5 Relative CCV
ALL	76.98	37.68	38.41	26.81	13.04
JWB	72.41	44.83	13.79	24.14	13.79
JIBS	81.08	29.73	54.05	37.84	10.81
JIM	84.21	42.11	57.89	21.05	21.05
MIR	84.21	47.37	36.84	10.53	15.79
IBR	68.57	32.35	32.35	29.41	8.82

Note: CCV1 for ALL is calculated based on n = 139; CCV2-5 are calculated based on n = 107.

of contextual control variables, we observe that JWB is lagging behind other IB journals in employing relative measures. We note that almost 45% of the studies in JWB are using only country dummies (CCV2), which is higher than the average (~38%) of all the IB journals. On the other hand, we note that the combined use of country dummies (CCV2) and absolute country-level measures (CCV3) for JWB was "only" less than 59%, far lower than the average for all IB journals (76%). Given the importance of contextual controls especially in cross-border IB research, it is concerning that such a large proportion of studies utilize crude controls such as country dummies or absolute CCVs that capture very little of the underlying substantive reason for potential confounding effects at the country level (Teagarden et al., 2018).

4. Recommendations and conclusion

If international business research wishes to increase its ability to inform other research areas through citations, the top journals in the field must strive for higher methodological standards. Three key areas are the selection, use, and reporting of control variables, which has hitherto received little attention in the IB literature. The purpose of our article was to (1) analyze the current treatment of controls in IB research, (2) compare it to other social science published research, and (3) provide specific recommendations to guide future authors, reviewers and editors involved with IB journals. Hence, based on our analysis of important dimensions related to controls in IB regression studies, we propose the following nine sequential recommendations regarding the selection, use, and reporting of control variables in IB econometric studies.

4.1. Recommendations regarding selection of CVs

4.1.1. Recommendation 1

As a minimum standard of practice (MSP), authors should justify the inclusion of each CV in the study and explain how and why it may exert a biasing rather than substantive influence in your model. This justification should be grounded in theory. Authors must also provide detailed information about measurement of all CVs in the study and ensure that they are measured in accordance with the theory used to justify their inclusion. Point to previous studies that have measured specific CVs in a similar (or different) way and discuss any implications. It is furthermore good standard of practice (GSP) to provide evidence via citations of use of similar controls in similar settings to verify and certify the inclusion of each CV in the study.

4.1.2. Rationale

In the design stage of a study, scholars must justify inclusion/exclusion of relevant CVs based on theory and empirical evidence (CD1). Barring convincing explanations for choice of controls, the credibility of a study can be cast in doubt as it may lead to relevant (substantive) variance being treated as error variance and/or omitted variable problems. Indiscriminant selection and use of statistical controls increases the risk of both Type I and Type II errors by partialling true variance from the relationships of interest (Spector et al., 2000). The issue of justifying choice of CVs and clarifying which independent variables are treated as controls is important pre data-collection in order to ensure you have the correct variables accessible for inclusion as controls. Moreover, detailed information regarding the measurement of all CVs (CD2) is necessary to evaluate a study's validity and reliability (Becker, 2005; Breaugh, 2006). Even small amounts of measurement error in control variables may led to significant biases in outcomes (Griliches, 1977). Particular attention must be paid to match between theory and measurement (Boyd, Gove, & Hitt, 2005). Scholars must first demonstrate that measures employed plausibly capture the underlying theoretical constructs (Lawrence, 1997). In addition, the theoretical and empirical levels of analysis must match to avoid levels-of-analysis ambiguity (Klein, Dansereau, & Hall, 1994; Nielsen, 2014). It is good practice to justify, verify and certify previous empirical *use* of chosen controls by citing other studies that used similar controls in similar settings (CD3). Here it is important to emphasize that inclusion based on isomorphism (Atinc et al., 2012; Bernerth & Aguinis, 2015) is not adequate justification; do not simply include controls because others studies have done so; however, consult prior empirical studies that have used similar controls *in similar settings* in order to provide evidence of the theoretical relevance of a given control variable. Justification by citation (post data-collection) is inappropriate. Also, be aware that simply including more controls does not equate rigor or even conservatism in terms of tests of hypotheses (Carlson & Wu, 2012; Spector & Brannick, 2011). We do not, however, agree with Becker et al.'s (2015) recommendation that 'when in doubt, leave them out!'; rather scholars should eliminate doubt by letting theory guide their choices.

4.1.3. IB implications

Consulting prior empirical studies that have used similar controls in similar national or cultural (contextual) settings in order to provide evidence of the theoretical relevance of a given control variable is of particular importance in IB studies. If a control variable has been used in a different contextual setting to the present study, its relevance, appropriateness and validity may be questionable. IB studies may indeed benefit from explicitly testing the extent to which various controls act similarly across different national, political, institutional, or cultural settings. Furthermore, for IB research, where multilevel phenomena are commonplace (Nielsen & Nielsen, 2010; Peterson, Arregle, & Martin, 2012), it is particularly pertinent to ensure match between levels of theory and measurement. Far too often, national level contextual variables, such as institutions or culture, are proxied with measures at different levels. Similarly, contextual variables are also commonly used to measure individual (managerial) behavior. Indeed, even Hofstede's (1980) cultural dimensions are potentially subject to the ecological fallacy (Robinson, 1950), which arises because associations between two variables at the macro level (or ecological level) may differ from associations between analogous variables measured at the micro (or individual) level. Hence, the national level cultural dimensions identified by Hofstede (e.g., individualism − collectivism) cannot be applied to the individual (managerial) level. Hofstede warned about this logic already in his 1980 seminal piece and IB researchers need to pay careful attention to the nature and match between levels of theory and measurement with regards to controls (as well as other variables).

4.2. Recommendations regarding the use of CVs

4.2.1. Recommendation 2

As a minimum standard of practice (MSP), researchers should consider the theoretical implications of specific control variables in relation to key relationships before collecting data. Specify and verify your theory and, where relevant, hypotheses in terms of controls,

including predicted directionality of any CV-dependent variable relationship.

4.2.1.1. Rationale. Theorizing about important control variables and/or explaining the dependent variable in terms of its residual ensures theoretical relevance and provides stronger evidence of boundary conditions of a theory and its predictions. Prediction of sign (positive or negative) for any CV-dependent variable relationship (CD4) is key to understanding potential impact of inclusion a priori (Becker, 2005; Carlson & Wu, 2012; Bernerth & Aguinis, 2015).

4.2.1.2. IB implications. IB research is particularly vulnerable (or subject) to changes in theoretical boundary conditions due to contextual variation (Teagarden et al., 2018). Such changes in boundary conditions may result in a theory's applicability being limited to particular contexts, for instance countries with (in)efficient market mechanisms or particular legal systems (Cuervo-Cazurra et al., 2016). IB researchers must pay particular attention to controls that potentially limit the applicability of the arguments to certain contexts and discuss theoretical and empirical implications of inclusion/exclusion, as well as prediction of directionality of sign, where appropriate.

4.2.2. Recommendation 3

As a minimum standard of practice (MSP), researchers should analyze correlations between CVs and all other variables (both independent and dependent variables) in the study. Calculate VIFs for all independent variables, including controls.

4.2.2.1. Rationale. Including controls with unknown, high correlations with other independent variables (CD6) may confound statistical control efforts, potentially disrupting efforts to purify relationships and further restricting residuals (Carlson & Wu, 2012; Spector & Brannick, 2011). It should be noted that as long as controls are not collinear with variables of interest, there is little concern regarding multicollinearity. Including controls that are not associated with dependent variable(s) (CD5) may reduce power to detect meaningful relationships. There may be strong theoretical reasons for including controls with little or no association with dependent variables, in which case authors are recommended to write something like "we also controlled for xyz, which have been identified in other studies (or theoretically) as relevant, however, as they were not significant and did not change the results, we omitted them from the final analysis". Note that while such statements may be published in the final version, reviewers should request that all relevant controls be included in earlier review rounds to avoid systematic capitalization on chance (Aguinis et al., 2017).

4.2.2.2. IB implications. IB studies are particularly subject to multicollinearity issues due to the need to control for contextual (i.e. home or host country) factors, many of which are highly correlated even if theoretically distinct. For example, Wan and Hoskisson (2003) analyzed firms from 16 European countries in terms of their factor endowments (natural, advanced and human) and political, legal and societal institutions, but had to group them into more or less munificent environments due to high correlations among theoretically different factors. By the same token, IB studies that analyze MNC subsidiary strategy and performance may find various HQ level controls (i.e. size, R&D intensity, or diversification) to be highly correlated with subsidiary level independent variables thus rendering inclusion problematic. Moreover, the issues of including controls with little or no association with the dependent variable(s) is also exacerbated in IB studies where sample size is often limited.

4.3. Recommendations regarding the reporting of CVs

4.3.1. Recommendation 4

It is considered good standard of practice (GSP) to write a separate section (for instance in the Methods section) dedicated to reporting on included control variables, their justification, measurement and relevance.

4.3.1.1. Rationale. A dedicated section that clearly outlines which variables are treated as controls, why they are included, and how they are measured (CD7) increases a study's validity and reliability as well as ensures replicability. Reporting on the treatment of controls aids both reviewers and readers and greatly increases the utility of the study.

4.3.1.2. IB implications. As noted earlier, authors in IB journals seemingly pay less attention to systematically reporting the nature of CVs than do authors in comparable management journals. More than 35% of studies across the five IB journals lack a dedicated section discussing included CVs, often leading these studies to provide inadequate information about criteria for inclusion, measurement, and potential influence on theoretical (and empirical) models. Reporting systematically on all controls in a dedicated section improves the reliability and replicability of IB studies, thereby increasing their potential for academic impact.

4.3.2. Recommendation 5

At a minimum (MSP), authors must report descriptive statistics for all continuous CVs in the correlation table, including means and standard deviations. As a good standard of practice (GSP), we strongly recommend also reporting ranges of CVs (and other variables). Report percentage of observations in each category for categorical (ordinal) variables where appropriate.

4.3.2.1. Rationale. Providing comprehensive descriptive statistics specifies and verifies the properties of the sample and variables. Means and standard deviations (CD8) provide important information about central tendencies and distributional attributes of the sample, whereas the range and minimum and maximum values (CD9) allow for detection of potential outliers.

4.3.2.2. IB implications. The nature of samples (and variables) can be very complex in IB research and this increases the importance of accurate reporting of descriptive statistics. In particular, we find many examples of IB research where the variability of the relationship between an independent and dependent variable is restricted (i.e. some values of either the independent or dependent variable are unlikely), for instance due to specific situational contexts such as strong/weak market regulation or institutional environments. In such instances of restricted variance interactions (see Cortina et al., 2015 for discussion and examples), it is particularly important to report descriptive statistics, such as scale, range, means and standard deviations for all variables, including controls.

4.3.3. Recommendation 6

It is considered minimum standard of practice (MSP) to report correlations between CVs and all other variables (both independent and dependent variables) in the study and highlight significant correlations. Avoid "included" statements of controls (i.e. industry controls) unless strictly necessary (i.e. due to an excessively large number) and still highlight if significant. Also report VIFs for all independent variables, including controls.

4.3.3.1. Rationale. Reporting correlations between CVs and other variables verify the nature and relevance of CVs (CD10). All included variables in a study are important and correlations among these

variables provide important information about the data and model utility.

4.3.3.2. IB implications. Neglecting to include one or more controls in the correlation table is always problematic, however, for IB studies this may be exacerbated if the CVs left out are related to the context of the study. For instance, controlling for contextual effects by country dummies without providing information about correlational relationships to IVs or DVs may hide important confounding information that influence the reliability of results.

4.3.4. Recommendation 7

As a minimum standard of practice (MSP), begin by reporting a base model including controls only in the regression table. We also recommend reporting a model which includes all independent variables without controls as good standard of practice (GSP).

4.3.4.1. Rationale. The base model including only controls (CD11) allows for incremental prediction as subsequent models provide evidence of utility of other independent variables (Carlson & Wu, 2012). The model with all independent variables but no controls (CD12) may provide additional information about the utility of IVs to explain un-controlled variation in dependent variable(s) (Becker et al., 2015; Breaugh, 2006, 2008).

4.3.4.2. IB implications. IB studies are particularly sensitive to omitted variable problems due to the complexity of spanning multiple environmental (i.e. political, economic, socio-cultural, and institutional) contexts. Combined with multifaceted, multilevel relationships (such as performance implications of international diversification, see Bausch & Krist, 2007; Nielsen & Nielsen, 2010) and often less-than desirable sample sizes (i.e. relatively low power), this puts a premium on illustrating the utility of a particular model. Entering contextual CVs first may lead to under-estimation (misattribution) of variance that should be attributed to conceptually important IVs at lower (i.e. firm) levels (Goldstein, 2011; Estrin, Nielsen, & Nielsen, 2017).

4.3.5. Recommendation 8

Authors must (MSP) report the influence of controls, whether significant or not, on their findings. We recommend that researchers report effect sizes in future submissions to IB journals. Dedicate a specific paragraph of the results section of the paper to discussion of controls.

4.3.5.1. Rationale. Information about the influence of controls on study findings (CD13) is essential for interpretation of reliability, generalizability and replicability of results. Interpreting final results of model testing both with and without controls (see recommendation 7 above) is essential to ruling out controls as potential explanation for study results. If and when results differ between models with and without controls, these differences must be discussed and the study findings interpreted accordingly (Becker, 2005). Statistical significance is a function of sample size, effect size, and threshold value of p – not a direct indicator of size of effect. Effect size provides important information about practical interpretation, generalizability and facilitates comparisons across studies (Bettis, Ethiraj, Gambardella, Helfat, & Mitchell, 2015).

4.3.5.2. IB implications. Given the importance of context in IB studies, final results must be interpreted carefully in terms of controls in order to ensure generalizability beyond the specific context of the sample. Moreover, it is highly recommended that IB scholars consider using research design in place of control variables to address various biases. For instance, researchers may try to find natural experimental contexts or control for unmeasured effects through careful sample stratification. We also note that 98% of the IB studies in our sample did not report effect sizes for any associations in their results sections – a prerequisite in an increasing number of top academic journals nowadays.

4.4. Recommendations regarding contextual CVs in IB studies

4.4.1. Recommendation 9

Researchers should include contextual control variables (CCV) in multi-country studies. Depending on the nature of the study and the importance of such contexts, we recommend usage of a hierarchy of CCVs, beginning with the simplest inclusion of country dummies via specific (absolute) home/host country variables toward relative measures of distance or difference between two or more countries. Results should be interpreted in terms of these contextual control variables.

4.4.1.1. Rationale. Any study that spans two or more country contexts should (at least) seek to control for potential confounding factors at the country level (CCV1) (Cho & Padmanabhan, 2005). While inclusion of country dummies (CCV2) may suffice in some situations, such an approach leaves much to be desired in terms of interpretation as such variables provide no substantive information about the country. Absolute CCVs (CCV3), such as GDP, GDP per capita, unemployment rate, or institutional quality in either home or host country provide more concrete information about the nature of the specific country context. Often, however, only one such measure can be included due to multicollinearity issues as mentioned earlier. Finally, controlling for aspects of differences between two or more country contexts by use of relative CCVs (CCV5), such as the Kogut and Singh (1988) cultural distance index, may be important in multi-country studies where the extent of relative difference is likely to influence (hinder) the independent-dependent variable relationship.

4.4.1.2. IB implications. Contextual differences become even more salient when applying extant theories and empirical findings across countries. IB research hinges on its ability to delineate contextual differences and explain how they influence transactions among actors as they cross national boundaries (Zaheer et al., 2012). Multi-country IB studies should (at a minimum) seek to control for some of these contextual influences in order to establish boundary conditions around the theory and empirical observations. While there may be instances even in IB studies where country dummies suffice, we strongly recommend inclusion of specific country level CCVs to capture the complexity of variation in such contexts. Multilevel IB studies may help tease out the specific sources of variation at various levels, including regional, country, industry and firm levels (Nielsen & Nielsen, 2010). Moreover, while relative distance measures, most commonly based on Hofstede's measures of national culture, have been used frequently in IB studies, such measures suffer from a number of shortcomings pertaining to their conceptual and methodological properties (Shenkar, 2012). While this is not the place to further discuss such issues, we do point to the importance of considering multi-directional distance measures (Schaffer & Riordan, 2003) that is measures that take into account the fact that (perceived) distance may differ depending on which country it is measured from. Related, we also feel a need to point out that intra-cultural variation often is stronger than inter-cultural variation and such effects are not captured by traditional distance measures.

By following these recommendations, IB scholars can improve their use and reporting of controls in empirical regression studies. Table 4 below offers an overview of the different control dimensions and their validation purposes and sequence. It also includes a checklist of recommendations according to Minimum Standard of Practice (MSP) and Good Standard of Practice (GSP), which may serve as a guideline for reviewers and editors to gauge the extent to which minimum requirements are met with regards to treatment of controls so as to warrant publication. Authors, however, are strongly encouraged to incorporate all dimensions of controls in order to justify, specify, verify and certify

Table 4
Recommendations.

Criterion	CD	#	Recommendation	Area of Validation	Sequence	MSP	GSP
Selection	CD1	**1**	Justify inclusion, measurement, and how and why each CV may exert a biasing rather than	Justify	Pre data-collection	✓	
	CD2		substantive influence in your model	Specify/ Verify	Pre data-collection	✓	
	CD3			Justify/ Verify/ Certify	Pre data- collection		★
Use	CD4	**2**	Specify and verify your theory and, where relevant, hypotheses in terms of controls, including predicted directionality of any CV-dependent variable relationship	Justify/ Specify	Pre data-collection	✓	
	CD5	**3**	Analyze correlations between CVs and all other variables (both independent and dependent	Verify	Post data-collection	✓	
	CD6		variables) in the study; calculate VIFs for all independent variables	Verify	Post data-collection	✓	
Reporting	CD7	**4**	Write a separate section dedicated to reporting on included control variables, their justification, measurement and relevance	Specify	Post data-collection		★
	CD8	**5**	Report means, standard deviations and ranges (and percentage of observations in each category for categorical (ordinal) variables) for CVs in correlation table	Specify/ Verify	Post data-collection	✓	
	CD9			Specify/ Verify	Post data-collection		★
	CD10	**6**	Report correlations between CVs and all other variables (both independent and dependent variables) in the study; report VIFs for all independent variables	Verify	Post data-collection	✓	
	CD11	**7**	First, report a base model including controls only; then report a final model including all	Verify	Post data-collection	✓	
	CD12		independent variables without controls	Verify	Post data-collection		★
	CD13	**8**	Report the influence of controls (including effect sizes) on your findings	Interpret	Post data-collection	✓	
Context	CCV1	**9**	Include contextual CVs (CCV) in multi-country studies and interpret results accordingly	Specify/ Verify/ Interpret	Pre- and post data-collection	✓	
	CCV2			Specify/	Pre- and post data-		★
	CCV3			Verify/	collection		
	CCV4			Interpret			
	CCV5						

*MSP: Minimum Standard of Practice; GSP: Good Standard of Practice.

Table 5
Standard of Practice Comparison Across International Business Journals.

Standard of Practice	Year	*Journal% (f)*					
		JWB	*JIBS*	*JIM*	*MIR*	*IBR*	*Average Across Journals*
Minimum > = 50%	***2015***	*53 (8)*	*67 (6)*	*80 (4)*	*80 (8)*	*67 (14)*	***69***
	2014	*100 (14)*	*78 (7)*	*83 (10)*	*100 (8)*	*84 (21)*	***89***
	2013	*94 (15)*	*79 (11)*	*86 (6)*	*92 (11)*	*89 (17)*	***88***
	2012	*78 (7)*	*92 (11)*	*80 (8)*	*80 (4)*	*93 (13)*	***85***
Average Across Years		*81*	*79*	*82*	*88*	*83*	***83***
Minimum > = 75%	***2015***	*13 (2)*	*11 (1)*	*60 (3)*	*30 (3)*	*14 (3)*	***26***
	2014	*43 (6)*	*33 (3)*	*33 (4)*	*75 (6)*	*32 (8)*	***43***
	2013	*50 (8)*	*29 (4)*	*71 (5)*	*58 (7)*	*47 (9)*	***51***
	2012	*11 (1)*	*42 (5)*	*10 (1)*	*40 (2)*	*57 (8)*	***32***
Average Across Years		*29*	*29*	*44*	*51*	*38*	***38***
GOOD > = 50%	***2015***	*13 (2)*	*33 (3)*	*80 (4)*	*20 (2)*	*19 (4)*	***33***
	2014	*36 (5)*	*22 (2)*	*50 (6)*	*50 (4)*	*48 (12)*	***41***
	2013	*44 (7)*	*50 (7)*	*86 (6)*	*33 (4)*	*58 (11)*	***54***
	2012	*11 (1)*	*50 (6)*	*30 (3)*	*80 (4)*	*43 (6)*	***43***
Average Across Years		*26*	*39*	*62*	*46*	*42*	***43***

f = Frequency; Note: Percentages are rounded to closest integer, n = 246.

the appropriate incorporation of controls in their regression studies.[5]

We close this perspectives article by re-examining our data on the use and reporting of control variables across the five IB journals in terms of Minimum Standard of Practice (MSP) and Good Standard of Practice (GSP). Table 5 below shows the percentage (and frequencies) of papers in each journal over the period that meet various thresholds with regards to MSP and GSP.

We start by evaluating the papers that meet 50% or better of the 9 MSP criteria (excluding CCV1[6]). The good news here is that close to 85% of the papers over the four years in each journal meets at least 50% of the MSP criteria and the trend is rather stable for all journals. However, looking at the same calculations for the 75% threshold of MSP, the picture is much bleaker. For all journals, below half of the papers meet this threshold over the four-year period. While MIR 'excels' at 51% on average, only 30% of their papers in 2015 met this threshold.

[5] Justifying and specifying are about adequately describing the selection and use of controls, whereas verifying and certifying are related to the empirical treatment and reporting of controls in regression studies. Together with interpreting, they all form the basis for appropriately selecting, using and reporting.

[6] Although inclusion of country-level controls is highly preferable in IB studies, not all IB studies are multi-country in nature. Thus we exclude CCV1 in MSP.

JIM has experienced the most variance at this threshold, while JIBS appears to be more stable, though only well below a third of their papers meet the 75% threshold. Finally, on average across the four years, less than half (43%) of the papers published in the five IB journals meet the 50% threshold of the five Good Standard Practice (GSP) criteria. In 2015 only 33% of the papers met the 50% threshold; more than half of the papers in JIM journal met this threshold. With the exception of IBR, there is relatively high variance across the four years.

Disheartening as these results may be, we find some comfort in the fact that IB research, on balance, does not seem to be worse (or better) than other comparable disciplines. Hence, the IB discipline seems to "remain at par with the methodological standards in adjacent fields for validity, reliability, replicability and generalizability" (Verbeke, Von Glinow, & Luo, 2017: 6) at least with regards to treatment of control variables. We also note that almost 85% of studies across the five journals, on average, met at least half of the nine Minimum Standard of Practice (MSP) criteria. On the other hand, no single study was found to have implemented all 9 of these MSPs. We appreciate that different journals may have different expectations and standards of reporting, as well as different space constraints, rendering it potentially difficult for authors to include all the Good Standard of Practice Recommendations provided here. Nevertheless, such variance in expectations and practices does not free researchers from the responsibility of scientific rigor when it comes to use and reporting of controls.

Together, these results illustrate the need for clearer guidelines with regards to controls in order to ensure transparency regarding the choices, measurements and procedures pertaining to the handling and reporting of statistical controls in econometric empirical studies. This perspectives article provides such guidelines in the form of a checklist and specific recommendations designed to help editors, reviewers and authors to systematically identify, use and report controls in IB studies.

We end by bringing a plea to IB scholars in general, and editors in particular. While the reason(s) behind what some label questionable research practices (Banks et al., 2016; Riley, 1958) with regards to control variables may be many and varied, most may essentially be driven by our academic reward system. Given increasing pressure to publish in "listed" journals to satisfy promotion, tenure and accreditation, there is a bias toward studies that find statistically significant support for a priori theory at the cost of interesting and/or un-supported findings (Banks & O'Boyle, 2013; Locke, 2007). Moreover, the increasing preoccupation with the often elusive "theoretical contribution" of most (IB) top academic journals makes empirical replication studies difficult to publish. Yet, replication and validation of empirical research on different samples and in different contexts – with the appropriate use of controls − may yield very valuable insights even if the hypotheses are not supported. This is particularly relevant for IB studies, where contextual boundary conditions may enhance theory development (Teagarden et al., 2018). Hence, we urge reviewers and editors to broaden scientific discovery by focusing on interesting research questions and rigorous methodological applications over *theoretical novelness* and support for a priori theory.

References

Aguinis, H., et al. (2017). Science's reproducibility and replicability crisis: International business is not immune. *Journal of International Business Studies, 48*(6), 653–663.
Ahlstrom, D. (2015). From the editors: Publishing in the journal of world business. *Journal of World Business, 50*(2), 251–255.
Andersson, U., Cuervo-Cazurra, A., & Nielsen, B. B. (2014). From the editors: Explaining interaction effects within and across levels of analysis. *Journal of International Business Studies, 45*(9), 1063–1071.
Atinc, G., Simmering, M. J., & Kroll, M. J. (2012). Control variable use and reporting in macro and micro management research. *Organizational Research Methods, 15*(1), 57–74.
Banks, G. C., & O'Boyle, E. H. (2013). Why we need industrial–organizational psychology to fix industrial–organizational psychology. *Industrial and Organizational Psychology, 6*(3), 284–287.
Banks, G. C., O'Boyle, E. H., Jr, Pollack, J. M., White, C. D., Batchelor, J. H., Whelpley, C. E., ... Adkins, C. L. (2016). *Questions about questionable research practices in the field of management: A guest commentary*.
Bausch, A., & Krist, M. (2007). The effect of context-related moderators on the internationalization-performance relationship: Evidence from meta-analysis. *Management International Review, 47*(3), 319–347.
Becker, T. E., Atinc, G., Breaugh, J. A., Carlson, K. D., Edwards, J. R., & Spector, P. E. (2015). Statistical control in correlational studies: 10 essential recommendations for organizational researchers. *Journal of Organizational Behavior, 37*(2), 157–167.
Becker, T. E. (2005). Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations. *Organizational Research Methods, 8*(3), 274–289.
Bernerth, J. B., & Aguinis, H. (2015). A critical review and best-practice recommendations for control variable usage. *Personnel Psychology, 69*(1), 229–283.
Bettis, R. A., Ethiraj, S., Gambardella, A., Helfat, C., & Mitchell, W. (2015). Creating repeatable cumulative knowledge in strategic management: A call for a broad and deep conversation among authors, referees, and editors. *Strategic Management Journal, 37*(2), 257–261.
Boyd, B. K., Gove, S., & Hitt, M. A. (2005). Construct measurement in strategic management research: Illusion or reality? *Strategic Management Journal, 26*(3), 239–257.
Breaugh, J. A. (2006). Rethinking the control of nuisance variables in theory testing. *Journal of Business and Psychology, 20*(3), 429–443.
Breaugh, J. A. (2008). Important considerations in using statistical procedures to control for nuisance variables in non-experimental studies. *Human Resource Management Review, 18*(4), 282–293.
Buckley, P. J., Elia, S., & Kafouros, M. (2014). Acquisitions by emerging market multinationals: Implications for firm performance. *Journal of World Business, 49*(4), 61–632.
Caprar, D. V., Devinney, T. M., Kirkman, B. L., & Caligiuri, P. (2015). Conceptualizing and measuring culture in international business and management: From challenges to potential solutions. *Journal of International Business Studies, 46*(9), 1011–1027.
Carlson, K. D., & Wu, J. (2012). The illusion of statistical control: Control variable practice in management research. *Organizational Research Methods, 15*(3), 413–435.
Cho, K. R., & Padmanabhan, P. (2005). Revisiting the role of cultural distance in MNC's foreign ownership mode choice: The moderating effect of experience attributes. *International Business Review, 14*(3), 307–324.
Clarke, K. A. (2005). The phantom menace: Omitted variable bias in econometric research. *Conflict Management and Peace Science, 22*(4), 341–352.
Conger, A. J., & Jackson, D. N. (1972). Suppressor variables, prediction, and the interpretation of psychological relationships. *Educational and Psychological Measurement, 32*(3), 579–599.
Cortina, J. M., Köhler, T., & Nielsen, B. B. (2015). Restriction of variance interaction effects and their importance for international business research. *Journal of International Business Studies, 46*(8), 879–885.
Cuervo-Cazurra, A., Andersson, U., Brannen, M. Y., Nielsen, B. B., & Reuber, A. R. (2016). From the editors: Can I trust your findings? Ruling out alternative explanations in international business research. *Journal of International Business Studies, 47*(8), 881–897.
Duanmu, J. L. (2014). State-owned MNCs and host country expropriation risk: The role of home state soft power and economic gunboat diplomacy. *Journal of International Business Studies, 45*(8), 1044–1060.
Duriau, V. J., Reger, R. K., & Pfarrer, M. D. (2007). A content analysis of the content analysis literature in organization studies: Research themes, data sources, and methodological refinements. *Organizational Research Methods, 10*(1), 5–34.
Ertug, G., Cuypers, I. R., Noorderhaven, N. G., & Bensaou, B. M. (2013). Trust between international joint venture partners: Effects of home countries. *Journal of International Business Studies, 44*(3), 263–282.
Estrin, S., Meyer, K. E., Nielsen, B. B., & Nielsen, S. (2016). Home country institutions and the internationalization of state owned enterprises: A cross-country analysis. *Journal of World Business, 51*(2), 294–307.
Estrin, S., Nielsen, B. B., & Nielsen, S. (2017). Emerging market multinational companies and internationalization: The role of home country urbanization. *Journal of International Management, 23*(3), 326–339.
Goldberger, A. S. (1991). *A course in econometrics.* Harvard University Press.
Goldstein, H. (2011). *Multilevel statistical models, Vol. 922.* John Wiley & Sons.
Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica: Journal of the Econometric Society,* 1–22.
Hofstede, G. (1980). *Culture's consequences.* Beverly Hills, CA: Sage.
Javidan, M., House, R. J., Dorfman, P. W., Hanges, P. J., & De Luque, M. S. (2006). Conceptualizing and measuring cultures and their consequences: A comparative review of GLOBE's and Hofstede's approaches. *Journal of International Business Studies, 37*(6), 897–914.
Kingsley, A. F., Noordewier, T. G., & Bergh, R. G. V. (2017). Overstating and understating interaction results in international business research. *Journal of World Business, 52*(2), 286–295.
Kirkman, B. L., Lowe, K. B., & Gibson, C. B. (2006). A quarter century of culture's consequences. A review of the empirical research incorporating hofstede's cultural value framework. *Journal of International Business Studies, 37*(3), 285–320.
Kish, L. (1959). Some statistical problems in research design. *American Sociological Review, 24*(3), 328–338.
Klein, K. J., Dansereau, F., & Hall, R. J. (1994). Levels issues in theory development, data collection, and analysis. *Academy of Management Review, 19*(2), 195–229.
Kogut, B., & Singh, H. (1988). The effect of national culture on the choice of entry mode. *Journal of International Business Studies, 19*(3), 411–432.
Lawrence, B. S. (1997). Perspective-the black box of organizational demography. *Organization Science, 8*(1), 1–22.
Locke, E. A. (2007). The case for inductive theory building. *Journal of Management, 33*(6), 867–890.

Makino, S., Isobe, T., & Chan, C. M. (2004). Does country matter? *Strategic Management Journal, 25*(10), 1027–1043.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage.

Nielsen, B. B., & Nielsen, S. (2010). In T. DeVinney, T. Pedersen, & L. Tihanyi (Eds.). *Understanding the internationalization-Performance relationship: A multilevel approach* (pp. 527–557). Emerald: Advances in International Management.

Nielsen, B. B. (2014). Construct measurement in management research: The importance of match between levels of theory and measurement. *Journal of Business Research, 67*(3), 403–406.

Pehazur, L. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated handbook*. Hillsdale, NJ: Lawrence Erlbaum.

Peterson, M. F., Arregle, J. L., & Martin, X. (2012). Multilevel models in international business research. *Journal of International Business Studies, 43*(5), 451–457.

Reeb, D., Sakakibara, M., & Mahmood, I. P. (2012). From the editors: Endogeneity in international business research. *Journal of International Business Research, 43*(3), 211–218.

Riley, J. W. (1958). Proceedings of the thirteenth conference on public opinion research. *The Public Opinion Quarterly, 22*(2), 169–216.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review, 15*(3), 351–357.

Schaffer, B. S., & Riordan, C. M. (2003). A review of cross-cultural methodologies for organizational research: A best-practices approach. *Organizational Research Methods, 6*(2), 169–215.

Shenkar, O. (2012). Cultural distance revisited. *Journal of International Business Studies, 43*(1), 1–11.

Spector, P. E., & Brannick, M. T. (2011). Methodological urban legends: The misuse of statistical control variables. *Organizational Research Methods, 14*(2), 287–305.

Spector, P. E., Zapf, D., Chen, P. Y., & Frese, M. (2000). Why negative affectivity should not be controlled in job stress research: Don't throw out the baby with the bath water. *Journal of Organizational Behavior, 21*(1), 79–95.

Stone-Romero, E. F. (2009). Implications of research design options for the validity of inferences derived from organizational research. *The Sage Handbook of Organizational Research Methods,* 302–327.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Procedures and techniques for developing grounded theory*. Thousand Oaks, CA: Sage.

Teagarden, M. B., Von Glinow, M. A., & Mellahi, K. (2018). Contextualizing international business research: Enhancing rigor and relevance. *Journal of World Business, 53*(3), 303–306.

Tong, T. W., Alessandri, T. M., Reuer, J. J., & Chintakananda, A. (2008). How much does country matter? An analysis of firms' growth options. *Journal of International Business Studies, 39*(3), 387–405.

Verbeke, A., Von Glinow, M. Y., & Luo, Y. (2017). Becoming a great reviewer. *Journal of International Business Studies, 48*(1), 1–9.

Wan, W. P., & Hoskisson, R. E. (2003). Home country environments, corporate diversification strategies, and firm performance. *Academy of Management Journal, 46*(1), 27–45.

Welch, C., & Piekkari, R. (2017). How should we (not) judge the 'quality'of qualitative research? A re-assessment of current evaluative criteria in International Business. *Journal of World Business, 52*(5), 714–725.

Zaheer, S., Schomaker, M. S., & Nachum, L. (2012). Distance without direction: Restoring credibility to a much-loved construct. *Journal of International Business Studies, 43*(1), 18–27.