

# A Cross-classified Analysis of the International Diversification-performance Relationship

Nielsen, Sabina; Nielsen, Bo Bernhard; Andersson, Ulf

*Document Version*  
Final published version

*Published in:*  
Proceedings of the Seventy-sixth Annual Meeting of the Academy of Management

*DOI:*  
[10.5465/AMBPP.2016.71](https://doi.org/10.5465/AMBPP.2016.71)

*Publication date:*  
2016

*License*  
Unspecified

*Citation for published version (APA):*  
Nielsen, S., Nielsen, B. B., & Andersson, U. (2016). A Cross-classified Analysis of the International Diversification-performance Relationship. In J. Humphreys (Ed.), *Proceedings of the Seventy-sixth Annual Meeting of the Academy of Management* (pp. 1218-1222). [14755] Academy of Management. Academy of Management Proceedings <https://doi.org/10.5465/AMBPP.2016.71>

[Link to publication in CBS Research Portal](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

## Take down policy

If you believe that this document breaches copyright please contact us ([research.lib@cbs.dk](mailto:research.lib@cbs.dk)) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 11. Jun. 2023



# A Cross-classified Analysis of the International Diversification-performance Relationship

**Sabina Nielsen, Bo Bernhard Nielsen, and Ulf Andersson**

Article in proceedings (Final published version)

**CITE:** A Cross-classified Analysis of the International Diversification-performance Relationship. / Nielsen, Sabina; Nielsen, Bo Bernhard; Andersson, Ulf. Proceedings of the Seventy-sixth Annual Meeting of the Academy of Management. ed. / John Humphreys. Briar Cliff Manor, NY : Academy of Management, 2016. p. 1218-1222 14755.

DOI: [10.5465/AMBPP.2016.71](https://doi.org/10.5465/AMBPP.2016.71)

Uploaded to [Research@CBS](https://research.cbs.dk): January 2018

# **A CROSS-CLASSIFIED ANALYSIS OF THE INTERNATIONAL DIVERSIFICATION-PERFORMANCE RELATIONSHIP**

BO BERNHARD NIELSEN  
University of Sydney  
Sydney, NSW 2006 Australia

SABINA NIELSEN  
University of Sydney

## **ABSTRACT**

We demonstrate how scholars can use RCM to investigate the impact of covariates at multiple levels, while accounting for the cross-nested nature that characterize strategy phenomena. Neglecting to account for the nesting structure of the data leads to biased estimates of variance components and standard errors, resulting in erroneous conclusions.

## **INTRODUCTION**

We argue that in order to explain the performance implications of firm international diversification, it is pivotal to account for the simultaneous embeddedness of firm strategies in industry and country contexts (Wan et al., 2011). These contexts provide the situational opportunities and constraints within which strategies are developed and executed, which, in turn, affects the ability of firms to extract value from international diversification. By introducing a multilevel perspective, we propose a unified and comprehensive contextual framework that has the potential to resolve this controversy and contribute to further empirical and theoretical developments in the field.

Empirically, it is important to account for the nested structure of the data accurately as failure to do so creates statistical problems related to disaggregation, violating the assumptions of independence across observations and ignorance of the inter-class correlations that increase the risk of type I and type II errors. Although typical multilevel analysis focuses on hierarchically nested random coefficients, differently structured (cross-nested or cross-classified) coefficients can also be incorporated in multilevel models and this paper aims at illustrating this approach on one of the fundamental IB questions.

We chose the setting of firms diversifying internationally because the relationship between international diversification (ID) and firm performance (P) represents one of the most important research areas in strategy and international business. Despite a long history of studying the relationship between ID and P, little consensus has been reached and empirical findings remain inconsistent. We argue that in order to explain the performance implications of firm ID, it is pivotal to account for the simultaneous embeddedness of firm strategies in industry and country contexts. Relying on a longitudinal sample of the world's 2000 largest MNCs, we illustrate the sequence of steps involved in analyzing the influence on the M-P relationship exerted by firm, industry, and country level variables over time. Our findings show that failure to account for the true (crossed) nesting structure of the data may lead to 1) biased estimates of variance components; 2) inflated standard errors of main effects leading to higher risk of type II error for lower level (e.g., firm) variables; and 3) risk of type I error for higher level (e.g.,

industry and country) variables. The wider implications for IB research are to caution scholars to pay adequate attention to the nesting structure of data when conducting IB research.

## MULTILEVEL MODELING

### Multilevel Nesting Structures

Firms are nested within institutional structures at both country and industry level. For instance, firms originating from a particular country are exposed to the same or similar formal and informal institutional characteristics, such as legal and educational systems, factor endowments, and socio-cultural norms (North, 1990). Moreover, firms share industry level influences, such as level of competition, barriers to entry, growth etc., which give rise to similarities in term of the competitive pressures they experience (Porter, 1990). Yet, countries and industries represent *imperfect hierarchies* because lower level units (firms) simultaneously may belong to multiple higher level units (industries and countries). In such non-hierarchical models the data structure is cross-classified since each firm uniquely belongs to a combination of both industry and home country. Cross-classified structures differ from perfect hierarchical nesting since the higher level factors (industry and home country) are not hierarchically nested (they cross) as not all companies from a particular country compete in the same industry, nor do companies competing in a particular industry all originate from the same country.

In multilevel cross-classified models, each unit potentially belongs to any combination of levels of the different factors. Cross-nesting of levels in data structures may occur at any level and correctly determining the level at which factors are crossed is important both theoretically and statistically. For instance, Hough (2006) examined business segment performance using data with a cross-classified nested structure at the highest level where yearly performance was hierarchically nested within business segments, cross-classified by corporations and industry. Others have estimated models where cross-classification occurs at intermediate level.

### Multilevel Analysis

Random coefficients modeling (RCM) is a methodology for the analysis of data with nested sources of variability. In the analysis of such data it is informative to take account of variability associated with each level of nesting (Snijders & Bosker, 2011). The multilevel structure of the data can be a result of either the sampling procedure (stratified random sampling) or natural nesting of behavior of the phenomenon under investigation. Datasets with a nested structure that includes unexplained variability at each level of nesting are not adequately represented by the probability model of multiple regression analysis. Instead, a RCM, an extension of multiple regression to a model including nested random coefficients, is recommended (Hox, 2010; Snijders & Bosker, 2011).

In RCM it is important to pay due attention to the nested structure of the data in order to avoid drawing wrong conclusions about observed relationships. RCM offers three substantial advantages over traditional statistical models: (1) improved estimation of effects within each unit; (2) formulation and testing of hypotheses about cross-level effects; and (3) portioning the variance and co-variance components among levels. Cross-classified RCM is a special case of RCM that allows for the separation of effects arising from complex non-hierarchical nesting structures. The aim is to separate out the effects of the cross-cutting hierarchies (e.g., country and

industry) on the dependent variable (e.g., firm performance). This is particularly important if there is a degree of association between the crossed levels. The use of cross-nested structures, and the ‘building’ into a statistical model the effects of the crossed levels, helps overcome such problems and avoid model under-specification, which may affect statistical estimates of other effects in the model. This is true both for effects of explanatory variables as well as the random effects of factors (levels) in the model.

### **AN ILLUSTRATION OF CROSS-CLASSIFIED MULTILEVEL MODELING OF THE INTERNATIONAL DIVERSIFICATION-PERFORMANCE RELATIONSHIP**

We selected the world’s 2000 largest firms based on sales in 2005 in order to capture the firms that generate most value around the globe. The number of companies per country was unequal (with the US, Japan, the UK, and Germany representing most companies), yet basing the choice on firm size made the companies in our sample comparable across countries and industries. Country level data was obtained from the *Global Competitiveness Report* published by the World Economic Forum; industry data was collected from the *Thomson One Banker* database. Our final sample consists of 1982 firms, representing 43 different home countries and 237 different industries, over a 15 year period (1992-2006).

*Firm performance* was measured as return on assets (ROA) (Combs, Crook & Shook, 2005), facilitating comparison of results with previous diversification studies (e.g., Lu & Beamish, 2004; Wan & Hoskisson, 2003). *Degree of internationalization (DOI)* was measured as foreign sales to total sales ratio (FSTS), which is one the most common measures of firm internationalization (Kirca et al, 2011). To test for a non-linear relationship between firm ID and P, we created quadratic and cubic product terms of this variable. We use primary industry affiliation (three-digit primary SIC code) for our industry measures. *Industry R&D* was measured as the average R&D intensity in each industry over the period. *Industry internationalization* was measured as the average foreign sales to total sales within each of the industries. We included *home country market size* measured as the logarithm of the country’s population and *formal* and *informal* institutions, measured through multiple items from the World Competitiveness Report.

### **RESULTS**

The parameter estimates of the final model support the S-curve relationship between ID and P. Comparing these estimates across the different nesting structures, the S-curve is robust across four of the five models; the only exception is model 1d, which fails to accounts for industry and country variance (all tables are omitted from this version). Hence, neglecting to adequately model the contextual embeddedness of firms within their home countries and primary industries may lead researchers to erroneous conclusions regarding the relationship between ID and P. This may explain inconsistent empirical results reported in the literature, as most prior studies fail to account for country or industry influences. Our results corroborate warnings that ignoring the nested structure of data will result in a loss of statistical power for testing lower level variables; thereby increasing the likelihood of committing Type II errors.

At the industry level, industry R&D is significant in two of the five models (Models 1d and 3d); these are the models that do not account for nesting within industry. While industry internationalization remains significant across all five models, the standard errors are underestimated in Models 1d and 3d. At the same time, the associated regression coefficients are

overestimated, thus resulting in higher significance levels compared to the other three models. These results parallel earlier findings that standard error estimates appear to be underestimated when the cross-classified structure is ignored, thus impacting the statistical significance inferences associated with the fixed-effect parameter estimates (Meyer, 2004). Consequently, misspecifications of the cross-nested structure may lead to type I errors (false positive).

At the country level, there is also divergence in the results across models. Home country market size is only significant in the models which fail to account for home country environment. Similarly, formal institutions are only significant in Model 2d, with deflated standard errors across the two models that do not account for country nesting (Models 1d and 2d). Although informal institutions remain significant across all models, the standard errors are again underestimated in Models 1d and 2d. Taken together, these results illustrate the potential danger of committing type I error by omitting a crossed level in cross-classified data structures.

The fixed effect of time appears significant only in Models 1d and 2d thus providing further evidence that inaccurately modeling of time may lead to biased estimates in strategy research. To illustrate this point further and demonstrate the advantages of multilevel longitudinal models, we proceed to compare the final model (5d) with OLS, fixed effects and random effects models. OLS does not account for the time dimension of the data and associated autocorrelation. As evident from Table 4, the OLS regression coefficients for degree of internationalization (DOI) are all significant but suggesting a different functional form than the theorized S-curve. This shows the importance of accounting for time when studying such relationships. In addition, the high significance of all other parameters illustrates how analyzing longitudinal data with regular OLS regression techniques artificially inflates sample size and leads to risk of committing type I errors.

The Random effects model produces similar results to Model 1d (see table 3), where the relationship between ID and P is not significant (thus indicating risk of type II error). Note also that the significance levels of the other predictors are similar to those in Model 1d. This is because both models are essentially two-level models that account for nesting of yearly performance within firms only. The results of the Fixed effects model (which is the most conservative test) support the S-curve hypothesis. However, fixed effects specification does not allow for appropriate testing of the model (i.e., for predictors at higher levels of analyses).

## **DISCUSSION**

Overall, our study has several important implications for IB research. First, the results provide strong support for the vital importance of developing cross-nested, multilevel models in IB research that account for variance at each theoretical level of nesting while allowing for interactions among factor across different levels. IB research is particularly ripe for complex model building and testing of this kind due to its inherent multilevel and often cross-classified nature, and our research suggests that IB scholars should start by acknowledging the theoretical and empirical nesting of the phenomena under investigation. Next, scholars must seek to identify relevant explanatory variables at each level, as well as consider potential cross-level interactions. From a theoretical perspective, our results also point to the potential overestimation of firm level effects on the M-P relationship while other effects at higher levels may be under- or misestimated in non-multilevel, non-cross-nested models. This echo's a recent call in the literature for more attention to contextual variables, such as industry and institutional variables,

which may be accomplished by combining theoretical perspectives from different levels of theory (Peng, 2009).

Our study illustrates the salience of paying attention to complex nesting structures when designing and executing strategy research. Modeling of time and contextual embeddedness is particularly important to strategy research as firm performance increasingly is shaped by changing opportunities and constraints beyond the boundaries of the firm (Porter, 1990; Wan & Hoskisson, 2003). Our study provides support for the importance of developing longitudinal cross-nested multilevel models in strategy research that account for variance at each theoretical level of nesting. Taking a step-by-step approach, we demonstrate how IB scholars can model time, firm, industry, and country effects simultaneously in a longitudinal cross-classified multilevel model to better account for complex non-hierarchical nesting. Doing so is important both theoretically and empirically and may help resolve inconsistencies in prior results as well as provide basis for new research questions.

Our results point to the potential misinterpretation of effects at the higher (crossed) levels in models that fail to accurately account for cross-nesting. The biased estimates of variance components and regression coefficients illustrate how ignoring the complex nesting of cross-classified data can lead to different conclusions pertaining to the influence of firm, industry, and country level factors on performance over time. Consistent with recent variance decomposition studies (e.g., McGahan & Victor, 2010), we show that time, firm, industry, and home country, as well as country-industry interaction, account for significant variance in firm performance. In addition, we demonstrate that ignoring the cross-nesting and the interaction between the crossed levels may misattribute variance to the lower level (e.g., firm). From a managerial perspective, such misattribution can have dire consequences as managers focus their attention too much on firm resources while ignoring important external contingencies.

Furthermore, ignoring cross-classification of nesting structures leads to biased standard errors and inferences based on statistical significance (Goldstein, 2011). This is important because IB researchers and practitioners may conclude that certain factors (e.g. industry R&D or home country market size) drive performance of firms, whereas these factors exhibit no significant effects when cross-classification is modeled correctly. Accurately modeling cross-classification is not only a matter of empirics but has important implications for theory development and testing. It also has important implications in practice as managers must decide on which environmental cues to respond to among increasingly complex information. As illustrated, one key advantage of CCRCM is its ability to account for time. Misestimating the ID-P relationship due to inaccurate treatment of longitudinal data may be costly as managers seek to optimize their diversification strategies.

## **REFERENCES AVAILABLE FROM THE AUTHORS**

Copyright of Academy of Management Annual Meeting Proceedings is the property of Academy of Management and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.