Journal of International Business Studies, 1–24

www.jibs.net

How well do supranational regional grouping schemes fit international business research models?

Ricardo Flores[1],
Ruth V Aguilera[2,3],
Arash Mahdian[4] and
Paul M Vaaler[5]

[1]School of Management, Australian School of Business, University of New South Wales, Sydney, Australia; [2]Department of Business Administration, College of Business, University of Illinois at Urbana-Champaign, USA; [3]ESADE Business School, Barcelona, Spain; [4]Wolfram Research, Champaign, USA; [5]Department of Strategic Management and Entrepreneurship, Carlson School of Management, University of Minnesota, Minneapolis, USA

Correspondence:
R Flores, School of Management, Australian School of Business, University of New South Wales, West Wing 5th Floor, ASB Building, UNSW Sydney, Australia NSW 2052.
Tel: +61 (2) 9385-6722;
Fax: +61 (2) 9662-8531;
email: r.flores@unsw.edu.au

Received: 28 November 2011
Revised: 4 March 2013
Accepted: 6 March 2013
Online publication date: 23 May 2013

Abstract
International business (IB) research has long acknowledged the importance of supranational regional factors in building models to explain phenomena such as where multinational corporations (MNCs) choose to locate. Yet criteria for defining regions based on similar factors vary substantially, thus undermining consensus regarding which regional grouping schemes fit IB research models better. In response, we develop and empirically validate a theory of comparative regional scheme assessment for model-building purposes assuming that: (1) schemes can be classified based on their source of similarity; and (2) schemes within the same similarity class can be assessed for their structural coherence, based on group contiguity and compactness. Schemes with better structural coherence will also exhibit better fit with IB research models. We document support for our theory in comparative analyses of regional schemes used to explain where US-based MNCs locate operations around the world. Geography-, culture- and trade and investment-based schemes with better structural coherence exhibit better initial fit with MNC location models and less change in fit after modest scheme refinement using a simulated annealing optimization algorithm. Our approach provides criteria for comparing similar regional grouping schemes and identifying "best-in-class" schemes tailored to models of MNC location choice and other IB research models.
Journal of International Business Studies (2013), 1–24. doi:10.1057/jibs.2013.16

Keywords: theory–method intersection; multinational corporations (MNCs) and enterprises (MNEs); foreign direct investment; regional strategy or strategies; simulation

INTRODUCTION

This paper develops and applies a theoretically grounded approach for comparatively evaluating the structural coherence of supranational regional grouping schemes and identifying which schemes better fit models explaining important research phenomena in IB and related fields. Consider the phenomenon of multinational corporation (MNC) location choice. More than 25 years of research in international business (IB) and related fields has been devoted to studying regional factors and proposing regional grouping schemes to explain whether and how individual countries attract MNC activity. In the 1980s, Ohmae (1985) highlighted a "Triad" of countries in North America, Western Europe and Greater Japan attracting more MNC activity than other similarly situated countries, given intra-regional advantages related to compact transportation, communications and similar business practices. Since then, IB researchers have investigated the impact of regional trade and investment patterns (Rugman & Verbeke, 2004), cultural traits (Ronen & Shenkar, 1985), institutional hazards (Slangen & Beugelsdijk, 2010), political traditions (Vaaler, 2008) and other regional dimensions that significantly and substantially affect MNC location choice and activity in individual countries.

Other fields have made this connection. MNC location and activity figure in the study of regional trade agreements in political science (Simmons, Elkins, & Guzman, 2006), the study of legal traditions in law and finance (La Porta, López-de-Silanes, & Shleifer, 2008), and the study of regional distance and foreign direct investment (FDI) as part of broader gravity model research in international economics and economic geography (Beugelsdijk, McCann, & Mudambi, 2010; Zwinkels & Beugelsdijk, 2010).

Research on the importance of regional factors often carries with it implicit if not explicit criticism of an alternative perspective assuming that MNC strategy tends toward a global scope (Bartlett & Ghoshal, 1991; Ghoshal, 1987; Yip, 1992), to take advantage of eroding national barriers to trade and investment, decreasing costs of travel and communication, and a general "flattening" of the earth (Friedman, 2005; Fukuyama, 1992). "Regionalists" hold that global MNCs are rare. More often, their locations worldwide are limited to one or two geographic clusters (Rugman & Verbeke, 2004), more in line with a regional scope of operations taking advantage of "semi-globalizing" trends (Ghemawat, 2003). Researchers outside IB, particularly in economic geography, also acknowledge these competing views on MNC location and activity, with additional nuances about how regional factors affecting MNC strategy in one cluster of countries occasionally spread to other clusters, resembling a global trend (Poon, 1997).

In highlighting the importance of regional factors for understanding MNC strategy, IB scholars also take on obligations to define regional grouping schemes with sound theoretical grounding and empirical precision. As we demonstrate below, schemes may be sensitive to small refinements that substantially change their impact on model fit – that is, their ability to explain data patterns as part of a statistical equation. IB research on regions is not uniquely vulnerable to such criticisms. Geography researchers since at least the late 1970s have noted the inadequacy of regional definitions (Walter & Bernard, 1978). Research based on mere division of the world by continents leads less often to insight and more often to "myths" about ecological, political and economic similarities among countries (Lewis & Wigen, 1997). Agnew (1999: 92) criticizes regional grouping schemes used in various geography subfields as "saying more about the political-social position of the observer than the phenomena the regions purport to classify".

If regional factors matter for understanding fundamental IB research issues such as MNC location choice, then IB researchers need rigorous theory and methods for comparatively assessing their model fit. Failure to meet that need undermines the development of consensus on regional strategy concepts, constructs and measures. In this context, we see an opportunity to contribute theory, methods and evidence to help IB researchers compare similar regional grouping schemes and identify the ones better fitted to particular IB research questions. Our approach eschews examination of the *ex ante* basis of any proposed scheme for grouping countries into regional groups. We give IB researchers latitude to choose their own regional grouping scheme and implement relevant analyses – say, analyses aimed at explaining MNC location choice worldwide. Instead, we propose an *ex post* analytical approach that can be applied to assess the model fit of schemes after their intended use.

First, regional schemes used by researchers are grouped into a "congruence class" based on identification of their underlying source of similarity – for example, schemes where countries are grouped together based on geographic proximity. Second, schemes within a congruence class are subjected to analyses of their "structural coherence" – that is, the extent to which countries within a regional group in the scheme are similar for purposes of exploiting a valuable common resource of interest. In the case of MNC location, this could be closer geographic access to similar intra-region markets. Our theory of structural coherence includes dimensions of "congruence", "contiguity" and "compactness". Contiguity is the absolute intra-group distance to valuable common resources. Shorter absolute distance across a group is better. Compactness is the intra-group distance to common resources relative to the group's center. Less variance in distance from the group center – a group shape more like a circle – is better. Schemes with better average contiguity and compactness across groups are schemes with better structural coherence. They are preferred schemes compared with others defined by a similar class of factors – what we call a "congruence class". Compared with other schemes in the same congruence class, we predict that these preferred schemes will exhibit better initial fit with the IB research model of interest, and change less in that model fit after modest scheme refinement.

Third, we empirically investigate evidence related to our prediction about the better structural coherence of preferred regional grouping schemes. To do this, we incorporate schemes into an empirical model explaining variation in data related to an IB research issue of interest. We assess the initial fit of that model based on alternative fit indicators: R^2 (McFadden, 1974), the Akaike information criterion (AIC; Akaike, 1974); and inspection of select parameter estimates. We then iteratively reassess model fit after modest scheme refinement by creating subgroups and transferring or exchanging countries between the subgroups. The overall aim of scheme refinement is to understand how, if at all, initial model fit might be improved by, for example, increasing the R^2. Scheme refinement and search for improved model fit are guided by a simulated annealing algorithm originally proposed for strategy research by Fox, Srinivasan, and Vaaler (1997). In an environment with countless alternative subgroupings, this algorithm promises a more diligent search of alternatives compared with conventional algorithms used in simulation research. We use alternative indicators of model fit to assess support for our prediction that preferred schemes will exhibit better initial model fit, with less change in fit after modest scheme refinement. Our *ex post* approach thus facilitates comparison of similar regional schemes, and the identification of preferred schemes meriting more frequent use in specific IB research models.

We demonstrate our *ex post* approach in the context of MNC location models and seven regional grouping schemes used to improve model fit. We sort them into three different congruence classes: geography, culture, and trade and investment. We assess their structural coherence. We identify the preferred scheme in each congruence class. We predict that, compared with other schemes in the same congruence class, these preferred schemes will exhibit greater initial fit with an MNC location model, and exhibit less change in fit after modest refinement. To test that prediction, we reduce each scheme to a series of regional dummy variables used to explain MNC location choice in a sample of 100 US-based MNCs operating in 105 countries in 2000. We estimate the likelihood of MNC location in a country based on regional dummies for a given scheme alone, and then as part of a larger model of MNC location including additional MNC-, industry- and country-level controls. Consistent with prediction, we find that models of MNC location incorporating preferred schemes exhibit better initial fit, with less change in fit after scheme refinement using simulated annealing.

Our study promises several contributions to research on regions and their impact on IB phenomena such as MNC location choice. First, we provide researchers with a novel *ex post* analytical approach for assessing the quality of regional grouping schemes, regardless of their *ex ante* grounding. Quality is based on comparing their initial fit and stability of fit after modest refinement. In this way, we promote the development of common scheme assessment standards and identification of "best-in-class" schemes for particular IB research issues. Second, and also related to our *ex post* analytical approach, we bring to IB research the first application of higher-order simulation and optimization techniques. Our simulated annealing search algorithm can search more diligently than conventional algorithms seeking to improve fit in models of MNC location choice. Third, the theoretical grounding of our *ex post* approach represents an advance in principles for assessing the structural coherence of schemes and predicting their model fit across a range of IB research contexts where countries are grouped by common geography, culture, trade and investment patterns, or other dimensions. The provenance of our theoretical framework in political science (Niemi, Groffman, Calucci, & Hofeller, 1990) and law and economics (Fryer & Holden, 2011) illustrates how theory development in IB can draw on related disciplines. Fourth, our empirical study documents support for several best-in-class schemes for improving fit on MNC location choice models where the schemes are based on geographic proximity (United Nations, 2007), cultural similarity (Ronen & Shenkar, 1985), or trade and investment commonalities (Donnenfeld, 2003). Our findings build on previous MNC location choice models (Flores & Aguilera, 2007). They also offer guidance to emerging research in IB and geography-related fields studying intra- and inter-regional distance and differences, and their impact on gravity models of bilateral trade and investment (Breschi & Lissoni, 2009; Fratianni, 2009; Fratianni & Oh, 2009; Singh, 2005). Fifth, the adoption of an *ex post* approach like ours helps integrate research on regions in IB with other geography-related fields. By highlighting the relativity of boundaries in regional grouping schemes, IB research moves closer to debates where space and place still matter, but description of the "fixed" physical environment increasingly includes more fluid symbolic, interactional, institutional, organizational and cognitive dimensions.

BACKGROUND CONCEPTS AND LITERATURE

The Regional Concept in IB Research

Definitions of "region" abound, but typically include reference to geographic areas, either part of a country or of the world, having definable characteristics but not necessarily fixed boundaries (DCE, 2012). Geography is fundamental to supranational regional definition in IB and related field research (Arregle, Miller, Hitt, & Beamish, 2013), with geographic proximity as a common assessment criterion for regional grouping schemes (Agnew, 1999). But geography is not the only basis for classifying regional schemes. IB research has identified schemes based on broad cultural dimensions (Ronen & Shenkar, 1985), trade and investment patterns (Donnenfeld, 2003), and other factors such as law and politics (Arregle et al., 2013). We review examples of such schemes, and note findings related to IB phenomena, particularly MNC location choice.

Regional Grouping Schemes Based on Geography

Geographic proximity is an intuitive organizing criterion for regional grouping schemes, with advantages related to ease of observation and measurement. Shared geographic borders between countries often imply other similarities. Indeed, geography-based regional grouping schemes often start with simple aggregation of countries based on shared continental borders. Kwok and Tadesse (2006), for example, use a continent-based regional grouping scheme to study the free-market orientation of financial systems in 41 countries. Similarly, Katrishen and Scordis (1998) use continents to determine MNC domicile and the economies of scale among MNC insurers. Geringer, Beamish, and da Costa (1989) also control for an MNC's continent of origin when assessing MNC performance effects linked to corporate diversification and internationalization levels.

Other geography-based regional grouping schemes in IB research subdivide continents to improve within-group proximity. Flores and Aguilera (2007), for example, use a 19-region scheme as part of a larger model of location choice for US-based MNCs in 1980 and 2000. On the other hand, geography-based groups could cross continents. Vaaler and McNamara (2004) combine Middle East with African countries as part of a seven-region scheme used to measure geographic market focus by major credit-rating agencies, and the impact of such focus on sovereign risk assessments published in the 1990s.

Regional Grouping Schemes Based on Trade and Investment

Regional grouping schemes based on trade and investment patterns are also prominent in IB research. We already noted Ohmae's (1985) Triad scheme comprising Greater Japan, North America and Western Europe (primarily France, Germany and the UK). He held that MNC survival required some dominant market positioning in at least one these regions. Building on Ohmae's insights, Rugman and Verbeke (2004) and others (Donnenfeld, 2003; Dunning, 2000) note that regional FDI by MNCs often follows multilateral trade relationships defined by Triad-related blocs such as the North American Free Trade Agreement (NAFTA) and the European Union (EU).

The underlying logic of trade and investment based regional grouping schemes is often economic similarity among members. Thus IB researchers define schemes based on country membership in international organizations defined by level of economic development: the Organization for Economic Cooperation and Development; the Organization for African Unity; or the Association of Southeast Asian Nations (Buckley & Ghauri, 2004; Gatignon & Kimberly, 2004). Related literature in political economy suggests that regional FDI follows more complex schemes based on multilateral regimes *and* bilateral trade and investment treaty arrangements (Simmons et al., 2006).

Links between bilateral country proximity and similarity, on the one hand, and bilateral patterns of trade and investment, on the other hand, have motivated a recent wave of IB research based on gravity models. These models have used continental membership and proximity to explain differences in MNC activity mix (Slangen & Beugelsdijk, 2010). Schemes based on multilateral agreements such as NAFTA reinforce an intrinsic attraction based on geographic proximity, and explain with gravity model assumptions additional trade and investment flows between countries within a regional bloc (Fratianni & Oh, 2009).

Regional Grouping Schemes Based on Culture

Other regional grouping schemes organize countries by broad cultural dimensions related to personal attitudes and beliefs. Perhaps the most prominent application of cultural dimensions to country groupings comes from Hofstede (2001), who surveyed IBM employees from 53 countries in the 1970s to derive several cultural dimensions of management behavior common to 12 regions.

As discussed by Kirkman, Lowe, and Gibson (2006), Hofstede's cultural dimensions and measures provided the basis for subsequent empirical studies by Kogut and Singh (1988) and others documenting similarity (dissimilarity) between MNC investment and competitive behaviors within (between) regions. Ronen and Shenkar (1985) developed their own regional grouping scheme of 45 countries in nine cultural clusters based on Hofstede. Furnham, Kirkcaldy, and Lynn (1994) devised yet another scheme of 41 countries in five cultural clusters, based on Hofstede. The GLOBE project produced other schemes based on Hofstede (House, Javidan, Hanges, & Dorfman, 2002). For example, Gupta, Hanges, and Dorfman (2002) used GLOBE project data in discriminant analyses to organize 61 countries into seven regional groups for purposes of understanding different patterns of individual and firm behavior. Cultural dimensions pioneered by Hofstede have also found their way into gravity models, with culture-based schemes explaining MNC FDI modes (Slangen, Beugelsdijk, & Hennart, 2011).

Regional Grouping Schemes Based on Other Factors

Other regional grouping schemes in IB research rely less on broad cultural indices and more on specific aspects of a country's culture. Perhaps best but still loosely described as institutional factors, their relevance to IB phenomena such as MNC location choice relates to how such factors alter basic rules of economic exchange (North, 1990; Xu & Shenkar, 2002).

In international finance research, Dyck, Volchkova, and Zingales (2008) use regional similarities in religion, media and political openness to explain differences in acquisition prices paid by investors to take control over firms. IB research by Vaaler (2011) uses regional groups to explain differences in venture capital availability across developing countries in the 2000s. In law and finance research, La Porta and colleagues have highlighted differences in the quality of government and international capital flows related to regions where certain legal systems dominate (La Porta et al., 2008). Regions such as Latin America, where legal systems from French civil law traditions dominate, also provide less protection to investors, thus limiting foreign investment compared with other regions where Anglo-American common law traditions dominate (Antràs, Desai, & Foley, 2007). On the other hand, comparative law research by Berkowitz, Pistor, and Richard (2003) documents variation in these trends, once more-refined regional groupings based on legal system are defined. Berkowitz and colleagues distinguish between countries and regions where common and civil law systems were imposed by force or developed organically. Countries where the legal system developed organically, whether civil or common law in nature, provide more protection than countries where the system was forcibly "imported".

Use, Validity and Reliability Issues in Research

Our summary review of different regional grouping schemes reveals clear differences regarding how and why researchers aggregate countries. No doubt some differences follow from research interests and experience. Other differences follow from theoretical perspectives and empirical methods thought appropriate for an issue under study. Even so, dimensions for grouping countries into regions sometimes lack any *ex ante* theoretical grounding, thus undermining concept, construct and measurement validity. Transcontinental groupings of countries from the Middle East and Africa used by Vaaler and McNamara (2004) follow from interviews with industry insiders – credit rating agency analysts – rather than from some theory of similarity within these regions. Alternatively, the purported theoretical grounding used to define some schemes has sparked debate over their validity. One example relates to the "Extended Triad" regional scheme used by Rugman and colleagues (Banalieva & Eddleston, 2011; Banalieva & Dhanaraj, 2013; Rugman, Li, & Oh, 2009; Rugman & Oh, 2012; Rugman & Verbeke, 2004). Critics have cast doubt on how similar countries are within the Extended Triad. For example, countries that might be sorted geographically into an "Africa and the Middle East" group instead constitute part of a "European" group. Countries arguably from an "Oceania" group fall instead into an "Asian" group (Asmussen, 2009; Clark & Knowles, 2003; Clark, Knowles, & Hodis, 2004; Osegowitsch & Sammartino, 2008).

Even where *ex ante* theoretical grounding is provided, we note cases where alternative regional grouping schemes based on similar theories and methods yield different results, thus impairing the reliability of constructs and measures. For example, Ronen and Shenkar (1985) refine Hofstede's regional clusters with results regarding MNC executive attitudes that differ from Hofstede's. Simmons et al. (2006) suggest that refinement of regional trading bloc definitions to account for bilateral investment treaties within blocs could change previous research findings about the impact of multilateral trade agreements on country imports and exports. Their claims matter for gravity-model-based research, which could yield different results regarding trade between countries within a regional trade bloc such as the EU, more narrowly defined by country signatories or more broadly construed to include signatory countries and other countries with trade bloc access through related bilateral agreements. These examples illustrate the potential for variance in the way that IB research theorizes about and then operationalizes regions to study important phenomena, including MNC location choice. No matter what the *ex ante* basis of these and other regional grouping schemes, reasonable refinements can lead to different results, with different implications for IB research and practice.

THEORY AND HYPOTHESIS DEVELOPMENT

Theoretical Framework Development

Ex post analytical approach

We take this concluding observation as a departure point for introducing our alternative analytical approach, and its theoretical grounding. Rather than attack the *ex ante* theoretical validity and reliability of prominent regional grouping schemes reliant on geography alone, or on geography and other factors, we propose an extended *ex post* analytical approach assuming the *prima facie* legitimacy of a scheme, but then empirically assessing its contribution to initial model fit and stability of fit after iterative scheme refinement and reanalysis. We contend that schemes contributing to better initial model fit and exhibiting less change after refinement tend to have greater research reliability and validity. These schemes explain important IB research phenomena such as MNC location choice with less concern about sensitivity to reasonable change, thus enhancing research reliability. Broader research validity is also enhanced to the extent that initial fit and fit stability after refinement are superior to others within a given scheme congruence class. They constitute better schemes within a class for studying important IB phenomena, and help "a community to collectively … ride upon common methods, schemas and templates" (Kogut, 2009: 711).

Structural coherence theoretical framework

We ground our *ex post* analytical approach in a broader theoretical framework of structural coherence informed by theory and evidence drawn from research in geography (Walter & Bernard, 1978), political science (Niemi et al., 1990), law and economics (Fryer & Holden, 2011), and IB (Arregle et al., 2013; Nachum, Zaheer, & Gross, 2008). The structural coherence concept motivating our framework comprises assessment of both group structure (that is, the geographic shape and size of groups within a regional grouping scheme) and scheme coherence (that is, the degree of within-group similarity and between-group dissimilarity). We propose that schemes with greater structural coherence will enhance the initial fit of models explaining IB phenomena such as MNC location choice, with less change in model fit after scheme refinement.

Three subconcepts relate closely to the structural coherence concept: structural congruence, contiguity and compactness. Structural congruence refers to criteria for classifying the source of similarity among group members – in our case, country members within a regional group. Geography scholars (Walter & Bernard, 1978) have long noted a lack of consensus about criteria for defining regions. Often, criteria identifying similarity among country group members are tailored to specific research or policy interests. For example, regional geographic groups are grouped by similar topography (e.g., the Andean countries of South America), natural resources (e.g., the Danubian countries of Europe), economic policies (e.g., the Central African CFA franc countries of Africa) or political-military alliances (e.g., the Southeast Asian Treaty Organization countries of Asia and North America). Our review of IB research on MNC location and activity suggests at least three broad classes of congruence: similar geography, typically based on shared country borders (Vaaler & McNamara, 2004; United Nations, 2007); and other classifications based on combinations of similar geography and broad cultural traits (Gupta et al., 2002; Ronen & Shenkar, 1985), or similar geography and trade and investment patterns (Donnenfeld, 2003; Rugman & Verbeke, 2004).

Within a given congruence class, we can identify multiple regional grouping schemes amenable to assessment of their structural coherency based on two other subconcepts: contiguity and compactness. Research in political science (Niemi et al., 1990) and law and economics (Fryer & Holden, 2011) has been central to the conceptual development and empirical measurement of group contiguity and compactness. Research leadership in these fields may follow in part from demands by policymakers in the US and other representative democracies for guidance on the geographic definition of legislative constituencies consistent with legal and/or constitutional mandates. From its earliest days as a republic, elected officials in the US sought electoral advantages through "gerrymandering" – that is, through redrawing geographic boundaries of constituencies so that the number of likely supporters would be increased, even if it would also lead to constituencies with irregular shapes. In the 20th century, US constitutional law has drawn on economics and political science research to help establish basic standards of geographic contiguity and compactness to limit gerrymandering.

Geographic contiguity refers to the basic connectedness of a group area. Within-group distance from end-to-end extremes constitutes one readily measureable indicator of contiguity. Increasing absolute geographic distance within a group renders access to common resources in the group more difficult and/or costly. In terms of MNC location choice, increasing absolute distance also increases costs of and difficulty in accessing common resources within the group such as markets regulated with a common legal system, communicating in a common language, unified by common customs and taxation treatment, and/or promoting common technological standards.

Geographic compactness refers to a group's geographic shape and what it means for distance to common resources relative to some referent point, usually the group's geographic center. Thus compactness is related to geographic shape and dispersion. Two measurable characteristics of compactness are perimeter length and variance in distance from the group center to perimeter points. A circular group shape minimizes center-to-perimeter-point variance, and improves compactness. Deviation from this ideal shape undermines individual group compactness, and thus relative access to common resources. In terms of MNC location choice, undermining compactness implies increasing relative costs of and difficulty in accessing common resources available within a regional group.

Hypothesis Development

Elements of our structural coherence framework – congruence class, contiguity and compactness – provide a basis for hypothesizing about the prospective contributions to model fit from different regional grouping schemes explaining MNC location choice. For a given scheme within the same congruence class – for example, a scheme based on broad cultural trait similarity – we can measure contiguity and compactness in each group, and calculate average contiguity and compactness for the scheme as a whole. These averages constitute measurable indicators of structural coherence within a congruence class. Lower average absolute distance across groups in a scheme indicates better group contiguity. More evenly dispersed distance relative to the center – essentially groups with circle shapes – indicates better group compactness. Previous IB research has documented MNC preferences for location near valuable common resources, thus making it easier to exploit regional markets and technologies (Nachum et al., 2008), or to adapt to regional law and politics (Arregle et al., 2013). Regional grouping schemes within a congruence class will provide such locational advantages to MNCs to the extent that these schemes are more structurally coherent – that is, when they have better average group contiguity and compactness. Compared with others in the same congruence class, these preferred schemes will increase overall fit in any model of MNC location choice. Thus we hypothesize that:

> **Hypothesis 1**: Models of MNC location choice based on preferred regional grouping schemes will exhibit better fit with less change in fit after scheme refinement compared with models based on other schemes in the same congruence class.

METHODOLOGY

Model Terms and Estimation

To evaluate evidence related to our hypothesis about the structural coherence of regional grouping schemes, we first define a model of MNC location choice incorporating various effects including those related to regional effects:

$$MNCSubsidiary_{ijlm} = \alpha_0 + \sum_{i=1}^{7} Country\text{-}LevelControls_i + \sum_{j=1}^{2} MNC\text{-}LevelControls_j + \sum_{l=1}^{16} IndustryDummies_l + \sum_{m=1}^{q} RegionalDummies_m + \varepsilon_{ijlm} \quad (1)$$

In Eq. (1), the dependent variable, *MNC Subsidiary*, is a 0–1 indicator, equal to 1 when US MNC j operating in industry l has a subsidiary in foreign country i grouped in region m. The likelihood of subsidiary location is explained by specific firm-, (MNC-) and country-level factors, as well as by 0–1 industry dummies and 0–1 regional dummies capturing unspecified but idiosyncratic industry- and region-related factors. All terms in Eq. (1) are fixed, except the number of regional dummies q, which varies with the type of scheme specified.

We define seven regional grouping schemes from three different congruence classes for inclusion in Eq. (1). Details regarding these schemes and the regional dummies they imply are given below. Comparison of model fit measures for different schemes within a given congruence class provides us with the basis for evaluating support for Hypothesis 1. If we constrain all non-regional dummy terms in Eq. (1) to zero then, estimations provide us with reduced model fit estimates of MNC location choice from a scheme. If we use all terms in Eq. (1), then we obtain model fit estimates of MNC location choice from the same scheme and other terms. We do both.

With the exception of regional dummies discussed below, variables, data and sampling follow substantially from previous research by Flores and Aguilera (2007), who modeled MNC location choice for 100 US-based MNCs in 1987 and 2000 to understand how MNC location drivers had changed in direction and magnitude over time. We use their MNC location data only for 2000 to remain consistent with the cross-sectional structure of Eq. (1). We use their *MNC Subsidiary* measure, which is a 0–1 variable taking the value 1 when a sampled MNC has a subsidiary located in country i. We use logistic regression to estimate location likelihood.

On the right-hand side of Eq. (1) we first include select variables from Flores and Aguilera (2007). We add variables familiar to gravity model research (Slangen & Beugelsdijk, 2010). The resulting two MNC- (*MNC-level Controls*) and seven country-level control (*Country-level Controls*) variables are (with predicted sign):

(1) *MNC ROI*, which is the MNC's return on investment (+);
(2) *MNC Size*, which is the number of employees in thousands for the MNC (+);
(3) host-country infrastructure (*HC Infrastr*), which is the total number of telephone lines per 1000 host-country residents (+);
(4) host-country wealth (*HC Wealth*), which is the natural log of host-country GDP in billions of US dollars (+);
(5) host-country size (*HC Size*), which is the natural log of host-country population in millions (+);
(6) common host-home country political system (*C Pol Syst*), which is a 0–1 dummy taking the value 1 when the host country is a democracy (+);
(7) common host-home legal system (*C Leg Syst*), which is a 0–1 dummy taking the value 1 when the host country has an Anglo-American common law legal system (+);
(8) common host-home language (*C Lang*), which is a 0–1 dummy taking the value 1 when the host country has English as its official language (+); and
(9) home-host-country distance (*H-H Dist*), which is the natural log of the geodesic distance between Washington, DC and the host-country national capital (−).

As additional controls, we include dummies for 18 of 19 four-digit North American Industrial Classification (NAIC) industry codes designated as the primary NAIC by a sampled MNC.

Non-regional Grouping Scheme Data and Sampling Sources

Data for these variables come from several sources. Data for the dependent variable, *MNC Subsidiary*, come from the *Directory of American Firms Operating in Foreign Countries* (Angel, 2001). This source defines a foreign subsidiary as a foreign entity with "substantial direct capital investment and [having] been identified by the parent firm as a wholly or partially owned subsidiary, affiliate or branch.

Franchises and non-commercial enterprises or institutions, such as hospitals, schools, etc., financed or operated by American philanthropic or religious organizations are not included" (Angel, 2001: i).

Data for our two MNC-level controls (*MNC ROI* and *MNC Size*) and industry dummies are from *Fortune* magazine's Fortune 500 list in 2000 (Fortune, 2001) and the *World Investment Report* of the United Nations Center for Transnational Corporations (UNCTAD, 2002). Data for our seven country-level controls include the World Bank's World Development Indicators (*HC Infrastr*, *HC Wealth*, *HC Size*), the *CIA World Factbook* (*C Pol Syst*, *C Lang*, *H-H Dist*), and Reynolds and Flores (1989) (*C Leg Syst*). Our final sample includes 9555 observations for 100 US-based MNCs operating in 105 countries in 2000.

Regional Grouping Schemes, Hypothesis Evaluation and Scheme Refinement Strategy

Regional grouping schemes

The number of regional dummies in Eq. (1) depends on which of seven regional schemes we use. Three schemes are based solely on geographic proximity:

(1) a seven-group scheme used by Vaaler and McNamara (2004);[1]

(2) a 19-group scheme used by the United Nations (UN, 2007);[2] and

(3) a five-group scheme using atlas-based continents ("Continents").[3]

Two more schemes combine geography with broad cultural traits related to Hofstede (2001):

(4) a 10-group scheme used by Ronen and Shenkar (1985);[4] and

(5) an 11-group scheme used by the GLOBE project researchers (Gupta et al., 2002) ("GLOBE").[5]

The last two schemes combine geography with trade and investment patterns:

(6) an eight-group scheme used by Donnenfeld (2003);[6] and

(7) a four-group scheme used by Rugman and Verbeke (2004).[7]

Hypothesis evaluation

Our *ex post* analytical approach includes comparison of regional grouping schemes within congruence classes defined by geographic proximity alone ("geography"), or geography and broad cultural traits ("culture"), or by geography and trade and investment patterns ("trade and investment"). With a congruence class, we assess structural coherence based on (lower) average group contiguity and (higher) average group compactness measures. Schemes within a congruence class with better measures are deemed to have better structural coherence. Accordingly, we calculate group contiguity and group compactness scores for the seven regional schemes. Group contiguity is measured as the longest point-to-point geodesic distance in miles within a regional scheme. Shorter distance indicates better group contiguity. Group compactness is measured as a quotient of regional area and perimeter length. Values range from 0 to 1. Values nearer to 1 – a perfect circle – indicate less difference in relative distance from the region's center, and thus better compactness.[8] Average group contiguity and compactness scores for each of the seven regional schemes in the three congruence classes appear in Table 1.

Note that one scheme in each class exhibits better contiguity and compactness scores. For the geography-based class, the UN (2007) scheme is preferred. For the culture-based class, the Ronen and Shenkar (1985) scheme is preferred, although differences from the GLOBE scheme are small. For the trade and investment-based class, the Donnenfeld (2003) scheme is preferred. Consistent with Hypothesis 1, we expect that these preferred schemes will be associated with models of MNC location choice

Table 1 Regional grouping scheme structural coherence characteristics by congruence classification

Scheme average	Scheme class						
	Geography			Culture		Trade and investment	
	Continents	Vaaler and McNamara (2004)	United Nations (2007)	R &D (1985)	GLOBE (2002)	Donnenfeld (2003)	Rugman and Verbeke (2004)
Contiguity	10,816	7785	4324	8462	8506	8462	13,922
Compactness	0.030	0.033	0.109	0.055	0.048	0.042	0.014

exhibiting better initial fit with less change after scheme refinement compared with other schemes in the same congruence class.

We use three indicators of model fit to evaluate support for Hypothesis 1:

(1) McFadden's adjusted pseudo-R^2 (MF-A Ps R^2) (McFadden, 1974);
(2) the AIC (Akaike, 1974); and
(3) inspection of the sign and significance of non-regional coefficients.

The MF-A Ps R^2 measures fit of MNC location choice models with higher measures for models using the same data and sampling indicating better fit (Long & Freese, 2006). MF-A Ps R^2 measure penalizes the addition of parameters if they do not add sufficiently to model fit.[9] This adjustment is critical. It compensates for preferred and alternative schemes having different numbers of groups initially. Our scheme refinement strategy described below involves the creation of new subgroups of countries. Fit measurement adjusted for the number of parameters used accommodates this refinement process. Consistent with Hypothesis 1, we expect that models based on preferred schemes will exhibit higher initial MF-A Ps R^2 measures with less change after refinement compared with alternative schemes in the same congruence class.

The AIC measure assesses model fit based on the concept of "information loss" from exclusion of additional parameters (Wagenmakers & Farrell, 2004).[10] Like MF-A Ps R^2 estimates, the AIC measure penalizes the addition of parameters not sufficiently reducing information loss. Unlike MF-A Ps R^2, the AIC measure itself has little meaning for fit assessment. It does have meaning when comparing a model with the lowest AIC measure in a given class with AIC measures for alternative models in the same class. Comparison of the model with the lowest score against others in the same class implies a comparison of the information-loss-minimizing model against others with some likelihood that they, too, may be information-loss minimizing. We can compare AIC measures for schemes within a given class and calculate comparative likelihoods of equivalent information-loss minimization to evaluate support for Hypothesis 1. Consistent with Hypothesis 1, we expect that models of MNC location choice based on preferred regional schemes will exhibit lower AIC measures with less change after refinement compared with alternative schemes in the same congruence class.

A third measure of model fit relates to non-regional grouping scheme dummy variables in Eq. (1). Direct measures of fit such as MF-A Ps R^2 and AIC may differ little between preferred and alternative schemes within a congruence class. Yet there will be indirect indicators that could differ substantially. Models of MNC location choice with regional grouping schemes exhibiting better structural coherence are also less likely to produce omitted-variable bias and more likely to yield precise estimates for other terms in Eq. (1). This implies more consistent signs at commonly accepted levels of significance for other MNC- and country-level controls. Thus, consistent with Hypothesis 1, we expect that models of MNC location choice based on preferred schemes will exhibit more MNC- and country-level controls with predicted sign and significance initially, and with less change after refinement, than the same controls estimated with alternative schemes in the same congruence class.

Scheme refinement strategy

Our *ex post* analytical approach depends on the ability to search diligently for refinements to original regional grouping schemes that could potentially improve model fit. Otherwise, prospective changes in scheme structure that might improve model fit could be left undiscovered, and confidence in evidence related to Hypothesis 1, the underlying structural coherence framework, and broader research aim of identifying "best-in-class" schemes will be undercut. We address that challenge first by designating a criterion for scheme refinement, minimization of model error sum of squares (ESS). A simple search criterion for refinement, ESS is essentially the squared difference between the average predicted and observed value of the dependent variable, *MNC Subsidiary*.[11] This contrasts with measures of model fit such as MF-A Ps R^2 and AIC. They are based on likelihood ratios adjusted for the number of parameters used to obtain that model fit. Our choice of ESS follows other research based on simple minimization criteria, including Fox et al. (1997), who use ESS to refine industry structure to explore the impact of intra-industry strategic group substructures affecting firm performance.

In concept, the number of countries and their combinations within a given group limit refinements potentially minimizing ESS. If the number is large, then it may be infeasible to search all possible alternative schemes in each group. A partial search seeking to refine the initial regional grouping scheme based on some simple optimization criterion may reduce search time, but other problems may arise. If the number of possible refinements is large, then simple minimization using conventional algorithms such as Newton–Raphson or Davidson–Fletcher–Powell could conclude refinement at a local ESS minimum. We may end the search prematurely, thus leaving the global ESS minimum unidentified, and researchers concerned that the refined scheme could be refined still further.

Refinement of regional grouping schemes based on a simulated annealing search algorithm improves on such shortcomings. It simulates real-world processes designed to reshape and strengthen metal by heating and then working it as slow cooling transforms the metal from a molten to a solid state. Annealing transforms the state of the atoms within the metal from energized agitation to a de-energized state. Early on, innate random variations allow slow-cooling atoms to escape local energy minima and find alternative, often denser structures. This strengthens the metal. In the context of an optimization algorithm applied to refinement of a regional grouping scheme, simulated annealing means that refinement paths can occasionally "escape" from local ESS minima and restart the search for even lower ESS values elsewhere. This property implies a more diligent search of scheme subgroups. Search based on simulated annealing does not ensure discovery of a global minimum, but it tends to do better than conventional search algorithms (Alrefaei & Andradóttir, 1999; Goffe, Ferrier, & Rogers, 1994).

Perhaps the best-known application of simulated annealing is to the "traveling salesman" problem, where the goal is to find the minimum trip distance connecting several cities. Academic applications of simulated-annealing-based search and refinement range from optimal land use and irrigation design (Aerts & Heuvelink, 2002) to microcircuit design (Kirkpatrick, Gellatt, & Vecchi, 1983). In management, aside from Fox et al. (1997), Han (1994) used simulated annealing to identify optimal information filing systems. Carley and Svoboda (1996) used this search algorithm to identify superior organizational adaptation strategies in the face of environmental shocks. Semmler and Gong (1996) used simulated annealing to identify optimal industry group size in analyses of real business cycle parameters.

We illustrate our application of regional grouping scheme refinement in Figure 1, which illustrates the *ex post* analytical approach. It is an iterative process, beginning with initial model analysis, notation of parameter estimates and indicators of model fit, and retention of initial model error sum of squares (ESS_{old}).

Define a model of MNC location choice based on terms in Eq. (1) and whichever scheme we seek to refine. The initial scheme, $P_{s\ old}$, is composed of u groups ($p_s=1, p_s=2, \ldots, p_s=u$), which are represented in Eq. (1) by $u-1$ 0–1 dummies. Logistic regression coefficients are then estimated. Next, a new partition, $P_{s\ new}$, is created by varying the group structure of countries in the scheme. That new partition arises from one of four types of changes chosen at random:

(1) Division: a random division of a group, p_s, into two subgroups, $p_{s,v}=1$ and $p_{s,v}=2$ subject to the requirement that each subgroup, $p_{s,v}$, has at least three countries as members.
(2) Transfer: after division into subgroups, a random transfer of one country from one to another subgroup.
(3) Exchange: after division into subgroups, a random exchange of two countries between two subgroups.
(4) Reunion: after division into subgroups, a random reunion of two subgroups.

With each new partition based on one of these changes, a new set of 0–1 regional dummies for Eq. (1) is generated, and then the model is re-estimated. We then compare the ESS for this new partition, ESS_{new}, with the previous ESS_{old}. If ESS_{new} is less than ESS_{old}, then the new partition is adopted, and the search algorithm moves "downhill" to ESS_{new}. If ESS_{new} is greater than or equal to ESS_{old}, then acceptance is random, and based on a criterion developed by Metropolis et al. (1953).[12] Under this criterion, $P_{s\ new}$ could be rejected. This option is intuitive, given that ESS_{new} is higher than ESS_{old} (not lower, as desirable). With this option, $P_{s\ old}$ remains, and the search algorithm tries another new partition. Here, the key innovation of simulated-annealing-based search and refinement comes into play by allowing the search to escape at times from one locality to find a new search vector. Here, $P_{s\ new}$ could be accepted, which would result in a non-intuitive "uphill" algorithmic move.

While a conventional minimization of ESS might end at a local minimum, a simulated-annealing-based search algorithm is able to escape many local ESS minima and find lower ESS values. Two factors decrease the likelihood of such non-intuitive moves

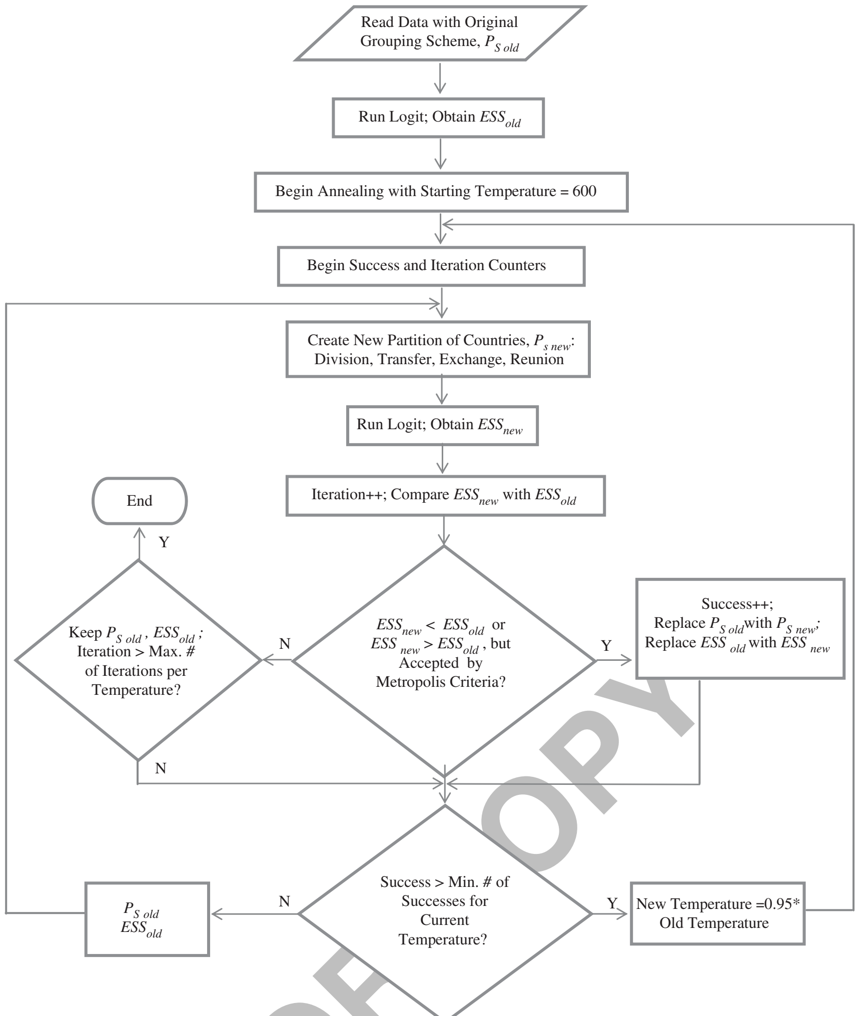

Figure 1 Simulated annealing optimization process.

based on the Metropolis criterion. A lower "temperature" in the annealing system makes it less likely to accept non-intuitive moves. At any one temperature in the annealing system, a larger non-intuitive move is also less likely.

After several attempts at new partitions, the starting temperature of the annealing system is reduced, and the annealing process continues. Typically, simulated annealing will include a "schedule" that specifies the initial temperature, the minimum number of attempts, and the minimum number of successful changes prior to taking a step and decreasing the temperature. As the temperature drops, larger non-intuitive moves uphill are discouraged, and the algorithm favors smaller refinements leading toward the global minimum. At the final temperature step in the schedule, simulated annealing concludes according to a stochastic stopping criterion based on a comparison of the rate of change in the system with computer speed in calculating such changes.[13] The annealing schedule – that is, the initial temperature, the number of iterations at each step, the number of successful perturbations to the system at each step, the size of stepwise temperature decreases, and the final stopping criteria – is *ad hoc*, and required experimentation. Our schedule is summarized in Table 2.

We implement this initial estimation and iterative refinement process featuring simulated annealing with a customized program written with Wolfram Mathematica Version 9 (Wolfram Research, 2012) and MATLAB Version 7.14 software (The MathWorks, 2012).[14] For each regional grouping scheme, execution of the estimation and iterative refinement program on a mainframe

Table 2 Simulated annealing schedule

Schedule parameter	Schedule value
Initial temperature	600
Reduction factor per temperature step	0.95
Minimum number of successes per temperature step	45
Maximum number of iterations per temperature step	50,000

computer takes approximately 10ëh, and averages from 90,000 to 100,000 iterations with 450–500 successful scheme refinements.[15] Schemes with fewer initial groups may have greater potential for change after refinement, but not necessarily greater model fit, based on indicators that adjust for parameter numbers, including the number of regional dummy parameters growing with refinement.

RESULTS

Results with Regional Dummies Only

To gain a preliminary understanding of how well regional grouping schemes fit models of MNC location choice, we first execute our program with regional dummies only. We exclude from Eq. (1) all other MNC- and country-level controls, as well as industry dummies. Select results are presented in Table 3. They include MF-A Ps R^2 and AIC measures, the number of regional dummies used and the number of such dummies with significant coefficients ($p < 0.10$), and the number of observations (9555). We report initial and refined figures for all Table 3 items.

These results offer preliminary support for Hypothesis 1. Of the three geography-based regional grouping schemes, the preferred UN (2007) scheme displays the highest (best) initial MF-A Ps R^2 estimate (0.165) and lowest (best) AIC goodness-of-fit estimate (6865). If we assume that the UN (2007) scheme is the most information-loss-minimizing scheme in its congruence class, then the likelihood that another geography-based scheme (e.g., Vaaler & McNamara, 2004) will yield similar information-loss minimization can be calculated as

$$\exp\left(\frac{\mathrm{AIC}_{\mathrm{UN2007}} - \mathrm{AIC}_{\mathrm{VM2004}}}{2}\right) = \exp\left(\frac{6{,}865 - 6{,}969}{2}\right) \approx 0 \qquad (2)$$

Table 3 Selected logistic regression results explaining MNC location choice (*MNC Subsidiary*) with regional dummies only

Items	Scheme class													
	Geography						Culture				Trade and investment			
	Vaaler and McNamara (2004)		United Nations (2007)		Continent (2012)		Ronen and Shenkar (1985)		GLOBE (2002)		Donnenfeld (2003)		Rugman and Verbeke (2004)	
	Initial	Refined	Initial	Refined	Initial	Refined	Initial	Refined	Initial	Refined	Initial	Refined	Initial	Refined
# Reg Dum	6	30	18	31	4	30	9	29	10	29	6	23	3	25
# Sig Reg Dum	5	27	16	27	2	22	6	27	9	26	4	18	3	23
MF-A Ps R^2	0.155	0.207	0.165	0.217	0.077	0.165	0.178	0.234	0.180	0.235	0.104	0.204	0.064	0.079
AIC	6969	6793	6865	6781	6999	6797	6921	6808	7033	6890	6948	6825	7067	6808
N	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555

Reg Dum is the number of regional dummies corresponding to each regional scheme (initial and refined). # Sig Reg Dum is the number of significant regional dummies corresponding to each regional scheme (initial and refined). Significance is deemed to be at 10% or higher levels. MF-A Ps R^2 is McFadden's adjusted pseudo-R^2 value. AIC is the Akaike information criterion value. N is the number of observations used for logistic regression estimation.

This likelihood is also nil when comparing initial AIC scores for United Nations (2007) and Continents schemes.

Changes in MF-A Ps R^2 and AIC scores after refinement are roughly the same or less for the preferred United Nations (2007) scheme compared with the Vaaler and McNamara (2004) and Continents schemes. Refinement of the United Nations (2007) scheme increases MF-A Ps R^2 by 0.052, less than the increase for the Continents scheme, and roughly the same increase as with the Vaaler and McNamara (2004) scheme (although from a lower initial starting point). The United Nations (2007) scheme AIC score decreases less than for either of the other two geography-based schemes. Such post-refinement comparisons are again consistent with Hypothesis 1. Models of MNC location choice based on preferred schemes with better structural coherence provide a better initial fit, with less change in fit after scheme refinement, compared with alternatives in the same congruence class.

Table 3 suggests similar support for Hypothesis 1 with the preferred culture-based scheme used by Ronen and Shenkar (1985). It exhibits higher initial MF-A Ps R^2 and lower AIC measures, with less post-refinement change in each, compared with the alternative culture-based GLOBE (2002) scheme. Similarly, we find support for Hypothesis 1 with the preferred trade and investment-based scheme used by Donnenfeld (2003). Compared with the alternative trade and investment-based scheme proposed by Rugman and Verbeke (2004), Donnenfeld (2003) exhibits higher initial MF-A Ps R^2 and lower initial AIC measures, with less post-refinement variation.

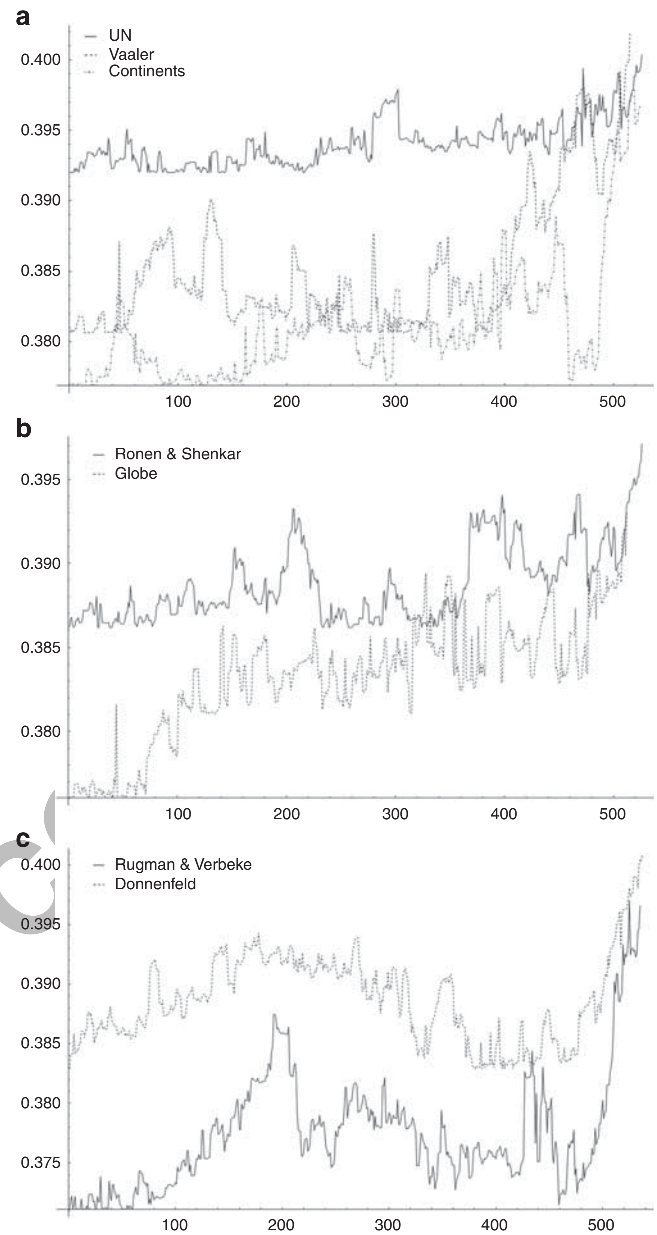

Figure 2 Regional grouping scheme refinement (successful change in ESS) and pseudo-R^2 estimates: (a) geography-based regional grouping schemes; (b) culture-based regional grouping schemes; (c) trade and investment-based regional grouping schemes.

Results with the Full Model

These preliminary results based on regional dummies only are largely confirmed by results obtained after execution of the program based on a full specification of Eq. (1), including regional dummies, industry dummies and all MNC- and country-level controls. Results are presented in Figure 2 and Table 4.

Figure 2 illustrates regional grouping scheme refinement trends based on our simulated annealing algorithm. The *x*-axes for all three graphs are the number of successful changes in group structure following either a proposed subgroup division, subgroup country transfer or exchange, or subgroup reunion. The *y*-axes for all three graphs are MF-A Ps R^2 measures. Initial measures are on the extreme left and fully refined measures are on the extreme right of each graph. Figure 2(a) graphs MF-A Ps R^2 values for all three geography-based schemes, whereas Figure 2(a) and 2(c) graph the same for the two culture-based and two trade and investment-based schemes.

Patterns of change in all three figures generally support Hypothesis 1. Preferred schemes start with higher MF-A Ps R^2 measures. In two of three

Table 4 Descriptive statistics and logistic regression results explaining MNC location choice (*MNC Subsidiary*) with full model (1)

Items	Mean (s.d.)	Scheme class													
		Geography						Culture				Trade & Investment			
		Vaaler and McNamara (2004)		United Nations (2007)		Continent (2012)		Ronen and Shenkar (1985)		GLOBE (2002)		Donnenfeld (2003)		Rugman and Verbeke (2004)	
		Initial	Refined	Initial	Refined	Initial	Refined	Initial	Refined	Initial	Refined	Initial	Refined	Initial	Refined
MNC ROI	14.2	0.01*	0.01*	0.01*	0.01*	0.01*	0.01*	0.01*	0.01*	0.01*	0.01*	0.01*	0.01*	0.01*	0.01*
	(8.7)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
MNC Size	6.7×10^{4}	9×10^{-6}***	9×10^{-6}***	9×10^{-6}***	9×10^{-6}*	9×10^{-6}***	9×10^{-6}***	9×10^{-6}***	9×10^{-6}***	9×10^{-6}***	9×10^{-6}***	9×10^{-6}***	9×10^{-6}***	9×10^{-6}***	9×10^{-6}***
	(6.5×10^{4})	(5×10^{-7})	(6×10^{-7})	(6×10^{-7})	(6×10^{-7})	(5×10^{-7})	(6×10^{-7})	(6×10^{-7})	(6×10^{-7})	(5×10^{-7})	(6×10^{-7})	(6×10^{-7})	(6×10^{-7})	(5×10^{-7})	(6×10^{-7})
HC Infrastr	234	-4×10^{-4}	-9×10^{-4}*	4×10^{-4}	-3×10^{-4}	2×10^{-4}	5×10^{-4}	6×10^{-4}	0.001†	9×10^{-5}	3×10^{-4}	-3×10^{-4}	−0.001*	-3×10^{-4}	-1×10^{-4}
	(221)	(4×10^{-4})	(4×10^{-7})	(5×10^{-4})	(6×10^{-4})	(4×10^{-4})	(4×10^{-4})	(5×10^{-4})	(5×10^{-4})	(3×10^{-4})	(5×10^{-4})	(4×10^{-4})	(5×10^{-4})	(4×10^{-4})	(4×10^{-4})
HC Wealth	24	0.94***	0.89***	0.95***	1.00***	0.98**	0.92***	0.82***	0.77***	0.93***	0.79***	1.05***	1.30***	1.10***	1.13***
	(2.0)	(0.06)	(0.07)	(0.07)	(0.08)	(0.06)	(0.07)	(0.07)	(0.08)	(0.06)	(0.07)	(0.06)	(0.08)	(0.05)	(0.07)
HC Size	16.4	−0.12*	0.002	−0.08	−0.11	−0.09†	0.02	0.003	−0.01	−0.15**	0.004	−0.23***	−0.35**	−0.22***	−0.15*
	(1.5)	(0.06)	(0.07)	(0.07)	(0.08)	(0.05)	(0.07)	(0.07)	(0.08)	(0.05)	(0.06)	(0.06)	(0.08)	(0.05)	(0.06)
C Pol Syst	0.65	−0.05	0.07	−0.11	−0.08	−0.06	0.14	0.22**	-1×10^{-5}	0.19*	0.16†	0.007	−0.34**	0.20**	0.15
	(0.48)	(0.08)	(0.10)	(0.09)	(0.10)	(0.08)	(0.10)	(0.08)	(0.09)	(0.08)	(0.09)	(0.08)	(0.10)	(0.07)	(0.10)
C Leg Syst	0.25	0.29*	0.33**	0.38**	0.87***	0.46***	0.95***	−0.09	−0.31†	0.13	0.21	−0.05	0.22	0.33**	0.15
	(0.43)	(0.12)	(0.16)	(0.13)	(0.16)	(0.12)	(0.17)	(0.12)	(0.16)	(0.11)	(0.14)	(0.12)	(0.16)	(0.12)	(0.14)
C Lang	0.19	0.36**	0.52**	0.36*	0.13	0.34**	−0.22	0.23†	0.26†	0.29*	0.35*	0.73***	0.38*	0.36**	0.22
	(0.39)	(0.13)	(0.17)	(0.15)	(0.16)	(0.13)	(0.16)	(0.14)	(0.15)	(0.14)	(0.15)	(0.13)	(0.16)	(0.12)	(0.14)
H-H Dist	8.5	−0.02	−0.24†	−0.43*	−0.36	0.42***	0.14	−0.46**	−0.55***	−0.19*	−0.21†	−0.49***	−0.52***	−0.05	−0.46**
	(0.47)	(0.13)	(0.14)	(0.21)	(0.23)	(0.11)	(0.14)	(0.09)	(0.11)	(0.09)	(0.11)	(0.09)	(0.10)	(0.10)	(0.14)
Ind Dum		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reg Dum		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant		−24***	−25***	−21***	−23***	−30***	−27***	−21***	−16***	−22***	−22***	−21***	−19***	−26***	−25***
MF-A Ps R^2		0.369	0.385	0.378	0.386	0.366	0.385	0.373	0.384	0.363	0.376	0.371	0.382	0.360	0.384
AIC		9328.2	8757.4	9218.3	8652.3	10192.6	9221.4	9082.6	8456.9	9059.9	8445.6	9898.9	8785.5	10336.3	10167.6
N		9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555

†$p < 0.10$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.
Abbreviations noted in Table 3 apply here. In addition, *MNC ROI* refers to MNC return on investment, *HC Infrastr* refers to Host-Country Infrastructure, *HC Wealth* refers to Host-Country Wealth, *HC Size* refers to Host-Country Size, *C Pol Syst* refers to Common Host-Home Country Political System, *C Leg Syst* refers to Common Host-Home Country Legal System, *C Lang* refers to Common Host-Home Language, and Ind Dum refers to industry dummies.

instances, these estimates exhibit little change until the end of the refinement process. The exception is with trade and investment-based schemes in Figure 2(c). The preferred Donnenfeld (2003) scheme starts with a higher initial estimate, but with substantial change from that initial measure after fewer than 100 successful refinements. Still, the Donnenfeld (2003) scheme exhibits less change in MF-A Ps R^2 measures compared with the same measures for Rugman and Verbeke (2004) after 100, 200 and 300 successful changes. Similarly, the preferred United Nations (2007) geography-based scheme and the preferred Ronen and Shenkar (1985) culture-based scheme both exhibit less change from initial MF-A Ps R^2 measures compared with the same measures for alternative schemes in the same congruence class after 100, 200, 300 and 400 refinements.

To confirm that the differences graphed in Figure 2 are not merely random, we repeat (but do not illustrate here) execution of the program 10 times for each of the seven regional grouping schemes. Observed patterns are consistent with those depicted in Figure 2. Models of MNC location choice based on preferred schemes exhibit better fit with fewer changes in other non-regional-group relationships after refinement.

Table 4 reports initial and refined MF-A Ps R^2 and AIC measures, as well as descriptive statistics and coefficient estimates for MNC- and country-level controls for program executions of fully specified versions of Eq. (1). We noted above initial and refined MF-A Ps R^2 measures indicating that models based on preferred United Nations (2007), Ronen and Shenkar (1985), and Donnenfeld (2003) schemes were superior to alternatives in the same congruence class. Evaluations of initial and refined AIC measures indicate the same contrasts, consistent with Hypothesis 1.

Table 4 also permits comparison of initial and refined coefficient estimates for MNC- and country-level controls. We find intuitive signs on most coefficient estimates before and after refinement. An MNC is more likely to have subsidiary operations in a given country if the MNC is larger and more profitable, and if the prospective host country is wealthier, less distant geographically, shares English with the US as its official language, and has a legal system based on Anglo-American common law traditions.

We compare these initial and refined coefficient estimates to observe differences in sign or significance across regional grouping schemes in the same congruence class. Models of MNC location choice based on preferred schemes will exhibit better consistency regarding MNC- and country-level signs and significance levels compared with coefficients based on other schemes in the same congruence class. This prediction also implies that models based on preferred schemes should exhibit less change in sign and significance after refinement.

Results here are mixed. For example, initial coefficient significance levels change for two terms in the preferred trade and investment schemed used by Donnenfeld (2003). An initially negative but not statistically significant sign on Host-Country Infrastructure (*HC Infrastr*) becomes significant at the 5% level after refinement. Similarly, an insignificant coefficient for Common Political System (*C Pol Syst*) turns significant (at the 1% level) and negative after refinement. Initial coefficient significance levels for four terms related to the alternative trade and investment scheme used by Rugman and Verbeke (2004) change after refinement: Common Host-Home Political System (*C Pol Syst*), Common Host-Home Legal System (*C Leg Syst*), Common Host-Home Language (*C Lang*), and Host-Home Country Distance (*H-H Dist*). The relative consistency of the preferred scheme used by Donnenfeld (2003) scheme supports Hypothesis 1.

We also find support for Hypothesis 1 when comparing coefficients before and after refinement for the two culture-based schemes. The initial coefficient significance level on one term, Common Host-Home Political System (*C Pol Syst*), changes after refinement of the preferred Ronen and Shenkar (1985) scheme, but two terms change in significance levels after refinement of the alternative culture-based scheme used in GLOBE (2002). They are Host-Country Size (*HC Size*) and Common Host-Home Political System (*C Pol Syst*). Again, the relative consistency of the preferred scheme, this time Ronen and Shenkar's (1985), supports Hypothesis 1.

For geography, however, the preferred UN (2007) scheme does not exhibit initial coefficient significance levels with greater resilience to scheme refinement. Three terms in the United Nations (2007) scheme shift in significance levels after refinement. They are Host-Home Common Legal System (*C Leg Syst*), Common Host-Home Language (*C Lang*) and Host-Home Country Distance (*H-H Dist*). Two of these terms (*C Lang*, *H-HDist*) also shift in significance level, while the third term (*C Leg Syst*) shifts substantially in coefficient magnitude when the Continents scheme is refined. Three terms shift in significance level for the Vaaler and McNamara (2004) scheme: Host-Home Common Legal System (*C Leg Syst*) again, plus two new terms – Host-Country Infrastructure (*HC Infrastr*) and Host-Country Size (*HC Size*).

Thus, for culture- as well as trade and investment-based schemes, all three measures of model fit – MF-A Ps R^2, AIC and non-regional dummy coefficients – support Hypothesis 1 and the underlying structural coherence framework from which Hypothesis 1 is derived. For geography-based schemes, two of these three assessment measures – MF-A Ps R^2 and AIC – support Hypothesis 1 and the underlying theoretical framework. We have noted above that the same pattern of results for MF-A Ps R^2 and AIC measures hold when MNC- and country-level controls as well as industry dummies are excluded.[16]

DISCUSSION AND CONCLUSION

Key Results and Contributions

Together, these findings constitute multiple bases of support for Hypothesis 1 and the theoretical framework from which it is derived. Initial model fit is better for certain regional grouping schemes within a given congruence class, whether that class be defined by geography alone (United Nations, 2007), by geography and other factors such as broad cultural traits (Ronen & Shenkar, 1985), or by trade and investment patterns (Donnenfeld, 2003). The stability of model fit is also generally better for those preferred schemes as they are refined through a diligent search of alternative subgroupings and related transfers and exchanges of subgroup countries. We document these findings when preferred and alternative schemes within a congruence class are the only factors explaining model fit, and when they are auxiliaries to other factors at MNC, industry and country levels. We uncover these findings with multiple indicators of model fit, nearly all of which point clearly at those same preferred schemes. At least with regard to regional factors and their impact on MNC location choice, we have a well-reasoned and broadly documented basis for identifying best-in-class schemes from among others that compete for IB research attention.

Our *ex post* analytical approach does not replace but complements *ex ante* approaches to defining schemes for IB research studying regional factors. We ground our complementary approach in a theory of structural coherence drawing on IB and related research in political science and law and economics. We enhance our complementary approach methodologically with a more diligent search and refinement algorithm to increase confidence in the findings it yields. We demonstrate how our complementary approach can be practically implemented in the context of MNC location choice models using alternative schemes to enhance model fit. Finally, we show how our complementary approach can identify best-in-class schemes meriting closer attention and perhaps more frequent use when studying the impact of regions and regional factors on MNC location choice.

Implications for IB Research and Practice

Our findings and contributions have implications for IB research and practice. We demonstrated how our *ex post* analytical approach could aid in evaluating different regional grouping schemes based on mere geography, or geography and broad cultural traits, or geography and trade and investment patterns. We could apply our *ex post* analytical approach to schemes based on other regional similarities. For example, we could apply it to schemes based on regional differences in legal system that might also attract or deter MNC location. We noted above relatively simple schemes based on coarse distinctions between Anglo-American common law and a few continental European civil law types proposed by La Porta et al. (2008). We also referred to more complex legal system schemes proposed by Berkowitz et al. (2003). With our *ex post* analytical approach, we could compare the initial and refined model fit characteristics of these schemes and others related to regional differences based on legal system. We could identify a best-in-class legal system scheme for explaining MNC location choice. We could broaden the scope of such study to other factors that might be used to organize regions: access to venture capital (Madhavan & Iriyama, 2009); technological standards (Clougherty & Grajek, 2008); and the prevalence of corruption (Wei, 2000) or bribery (Cuervo-Cazurra, 2008).

Expanding the scope of *ex post* analytical approaches like ours matters for the current debate between globalists and regionalists. Perhaps regionalist views would have greater cogency if regionalist advocates could point to regional grouping schemes and related findings that exhibit better initial fit and stability of fit after modest scheme refinement in models often used in IB research. There are similar benefits for gravity model research in IB and related fields such as economics and geography. Identifying and then using schemes that define regions with better initial fit and stability after refinement will lead to greater confidence in gravity-model-based evidence related to, say, the impact of regional agreements and networks on bilateral trade and FDI, as well as movements of skilled workers and the diffusion of knowledge and innovation (Breschi & Lissoni, 2009; Fratianni, 2009; Fratianni & Oh, 2009; Singh, 2005). Not only MNC research but also MNC management practice will benefit from identifying best-in-class, or at least better-in-class, schemes for thinking about where to locate subsidiary operations with a higher likelihood of benefiting from regionally proximate people, knowledge and markets.

Research debates in IB and geography gain from having theories, methods and evidence common to both fields. Nearly a decade ago, Sorenson and Baum (2003) mapped out a trajectory for geography research that defined the physical environment less as a fixed unidimensional resource and more as a flexibly defined resource with multiple dimensions. Space and place bring with them time-varying, symbolic, interactional, organizational and cognitive dimensions affecting local actors, including MNCs. We think our study, and the *ex post* analytical approach it highlights, helps foster that same kind of flexibility in "seeing" the space and place of regions in IB research. In this way, we contribute to closer coordination of research between IB and geography scholars.

Limitations and Future Research Directions

We think our study has several strengths. It also has limitations to be addressed in future research. We tout our *ex post* analytical approach as a means of classifying and then assessing regional grouping scheme structural coherence. We classified schemes based one or two dimensions readily apparent to us on observation. Yet schemes are more often multidimensional, thus rendering the classification step in our method more difficult to take. This limitation invites future research on means by which IB researchers might be able to simplify multidimensional schemes into one or two scheme factors, perhaps with the help of multivariate exploratory methods such as factor analysis.

Another limitation relates to the way we incorporate group contiguity and compactness in our analyses. We measure these two scheme attributes in advance of any empirical estimation and refinement. Of course, future research could take a different route by incorporating directly into those iterative estimations measures of average group contiguity and compactness, individually and in interaction with other regional attributes of research interest. This strategy promises deeper insight into how (and not just whether) scheme structural coherence matters for model fit and stability.

Another limitation relates to technology. We used simulated annealing to refine regional grouping schemes, but it required a customized program written for mainframe computer porting and execution. We are unaware of any such programs in standard statistical packages. On the other hand, conscientious researchers with a modicum of computer programming skills and assistance can develop an iterative refinement and reanalysis program capable of running on a mainframe or perhaps even a desktop or laptop computer. Our own program ran on a mainframe computer system, primarily because of the time and space required for our iterative logistic regressions. Other linear-based estimations do not impose the same time and space requirements.

For less programming-savvy researchers, our *ex post* analytical approach can still be partially implemented. Often there are only a few discrete refinements to an initial grouping scheme that are much more likely than others to generate changes in overall model fit. Such discrete refinement and reanalysis does not require advanced training in computer programming, or time on a mainframe computer. We are encouraged by the recent example set by Arregle et al. (2013). They study regional factors affecting the location choice of Japan-based MNCs. They test the robustness of their basic findings by simply replacing the original scheme and re-estimating with alternative schemes, including the UN-based scheme evaluated in this study. We would push them to demonstrate the robustness of their basic results to small refinements of the original scheme rather than substantially different schemes based on different congruence classes. Notwithstanding that push, we think their example merits commendation and imitation by others studying the significance of regional factors for MNC location choice and other IB research phenomena.

Our *ex post* analytical approach is well suited to examine regional factors tied to different regional grouping schemes in current use, and those now emerging in IB research. Our study analyzed some schemes (e.g., Donnenfeld, 2003) that might be rendered obsolete with, say, the next trade bloc dispute, disintegration and reformation. Our *ex post* analytical approach handles such eventualities. Researchers can go back into the class of remaining relevant schemes and help to identify another preferred scheme. We share Fawn's view (2009: 12) that, over time, there is little chance that any one definition of region will ever be "forced across researchers, even less so across disciplines …". We welcome continued innovation in scheme formulation, and the opportunity it presents to assess the structural coherence of new schemes compared with others in the same class. How well do regional grouping schemes fit IB research models? The answer co-evolves with innovation in scheme formulation and flexibility in scheme assessment.

ACKNOWLEDGEMENTS

We thank Randy Westgren and the University of Illinois at Urbana-Champaign Center for International Business Education and Research for financial support. We thank the University of Illinois at Urbana-Champaign Supercomputing Center, the University of Minnesota Office of Information Technology and Karl Smelker for technical support. Earlier drafts of this paper benefited from presentation at the Academy of International Business annual meeting, the Carlson School of Management and the Humphrey School of Public Administration at the University of Minnesota, the Henley Business School at Reading University, and the Fox School of Business at Temple University. Mark Casson, Isaac Fox, Martin Ganco, Robert Kudrle, Alan Rugman, Andrew Van de Ven, Richard Wang and Joel Waldfogel offered helpful comments, criticisms and suggestions for revision. We also thank to the co-editors of this special issue, Soerd Beugelsdijk and Ram Mudambi, as well as the three anonymous reviewers who helped us improve this manuscript. All errors are ours.

NOTES

[1]The geography-based regional grouping scheme adapted from the general scheme used by Vaaler and McNamara (2004) includes the following regions (in CAPITALS) and countries: AFRICA AND MIDDLE EAST: Algeria, Azerbaijan, Bahrain, Benin, Burkina Faso, Burundi, Democratic Republic of Congo, Egypt, Ethiopia, Gambia, Ghana, Guinea, Iran, Iraq, Israel, Ivory Coast, Jordan, Kenya, Kuwait, Lebanon, Liberia, Libya, Malawi, Mali, Morocco, Namibia, Niger, Nigeria, Oman, Pakistan, Qatar, Saudi Arabia, Senegal, Sierra Leone, South Africa, Sudan, Syria, Tanzania, Tunisia, Turkey, Uganda, United Arab Emirates, Yemen, Zambia; ASIA: China, Hong Kong, India, Indonesia, Japan, Malaysia, Philippines, Singapore, South Korea, Sri Lanka, Thailand; CENTRAL AND EASTERN EUROPE: Albania, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Russian Federation, Serbia and Montenegro, Slovenia; LATIN AMERICA: Argentina, Brazil, Chile, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Panama, Peru, Suriname, Uruguay, Venezuela; NORTH AMERICA AND CARIBBEAN: Canada, Dominican Republic, Jamaica, Mexico; OCEANIA: Australia, Fiji, New Zealand; WESTERN EUROPE: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom.

[2]The geography-based regional grouping scheme adapted from the full partition of worldwide regions in the world according to the information provided in the United Nations website includes the following regions (in CAPITALS) and countries: AUSTRALIA AND NEW ZEALAND: Australia, New Zealand; CARIBBEAN: Dominican Republic, Jamaica, CENTRAL AMERICA: Costa Rica, El Salvador, Guatemala, Mexico, Panama; EASTERN AFRICA: Burundi, Ethiopia, Kenya, Malawi, Tanzania, Uganda, Zambia; EASTERN ASIA: China, Hong Kong, Japan, South Korea; EASTERN EUROPE: Bulgaria, Czech Republic, Hungary, Poland, Romania, Russian Federation; Melanesia: Fiji; MIDDLE AFRICA: Democratic Republic of Congo; NORTHERN AFRICA: Algeria, Egypt, Libya, Morocco, Sudan, Tunisia, NORTHERN AMERICA: Canada; NORTHERN EUROPE: Denmark, Estonia, Finland, Ireland, Latvia, Lithuania, Norway, Sweden, United Kingdom; SOUTH AMERICA: Argentina, Brazil, Chile, Colombia, Ecuador, Peru, Suriname, Uruguay, Venezuela; SOUTH-EASTERN ASIA: Indonesia, Malaysia, Philippines, Singapore, Thailand; SOUTHERN AFRICA: Namibia, South Africa; SOUTHERN ASIA: India, Iran, Pakistan, Sri Lanka; SOUTHERN EUROPE: Albania, Croatia, Greece, Italy, Portugal, Serbia and Montenegro, Slovenia, Spain; WESTERN AFRICA: Benin, Burkina Faso, Gambia, Ghana, Guinea, Ivory Coast, Liberia, Mali, Niger, Nigeria, Senegal, Sierra Leone; WESTERN ASIA: Azerbaijan, Bahrain, Iraq, Israel, Jordan, Kuwait, Lebanon, Oman, Qatar, Saudi Arabia, Syria, Turkey, United Arab Emirates, Yemen; WESTERN EUROPE: Austria, Belgium, France, Germany, Luxembourg, Netherlands, Switzerland.

[3]The geography-based regional grouping scheme adapted from the scheme presented by United Nations website based on continents includes the following regions (in CAPITALS) and countries: AFRICA: Algeria, Benin, Burkina Faso, Burundi, Democratic

Republic of Congo, Egypt, Ethiopia, Gambia, Ghana, Guinea, Ivory Coast, Kenya, Liberia, Libya, Malawi, Mali, Morocco, Namibia, Niger, Nigeria, Senegal, Sierra Leone, South Africa, Sudan, Tanzania, Tunisia, Uganda, Zambia; AMERICAS: Argentina, Brazil, Canada, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Guatemala, Jamaica, Mexico, Panama, Peru, Suriname, Uruguay, Venezuela; ASIA: Azerbaijan, Bahrain, China, Hong Kong, India, Indonesia, Iran, Iraq, Israel, Japan, Jordan, Kuwait, Lebanon, Malaysia, Oman, Pakistan, Philippines, Qatar, Saudi Arabia, Singapore, South Korea, Sri Lanka, Syria, Thailand, Turkey, United Arab Emirates, Yemen; EUROPE: Albania, Austria, Belgium, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Norway, Poland, Portugal, Romania, Russian Federation, Serbia and Montenegro, Slovenia, Spain, Sweden, Switzerland, United Kingdom; OCEANIA: Australia, Fiji, New Zealand.

[4]The culture-based regional grouping scheme adapted from the general scheme developed by Ronen and Shenkar (1985) includes the following groupings (in CAPITALS) and countries: ANGLO: Australia, Canada, Ireland, New Zealand, South Africa, United Kingdom; ARAB, Bahrain, Kuwait, Oman, Saudi Arabia, United Arab Emirates; FAR EASTERN: Hong Kong, Indonesia, Malaysia, Philippines, Singapore, Thailand; GERMANIC: Austria, Germany, Switzerland; INDEPENDENT: Brazil, India, Israel, Japan; LATIN AMERICA: Argentina, Chile, Colombia, Mexico, Peru, Venezuela; LATIN EUROPEAN: Belgium, France, Italy, Portugal, Spain, NEAR EASTERN: Greece, Iran, Turkey; NORDIC: Denmark, Finland, Norway, Sweden; OUTSIDE (Countries not included in this scheme): Albania, Algeria, Azerbaijan, Benin, Bulgaria, Burkina Faso, Burundi, China, Costa Rica, Croatia, Czech Republic, Democratic Republic of Congo, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Ethiopia, Fiji, Gambia, Ghana, Guatemala, Guinea, Hungary, Iraq, Ivory Coast, Jamaica, Jordan, Kenya, Latvia, Lebanon, Liberia, Libya, Lithuania, Luxembourg, Malawi, Mali, Morocco, Namibia, Netherlands, Niger, Nigeria, Pakistan, Panama, Poland, Qatar, Romania, Russian Federation, Senegal, Serbia and Montenegro, Sierra Leone, Slovenia, South Korea, Sri Lanka, Sudan, Suriname, Syria, Tanzania, Tunisia, Uganda, Uruguay, Yemen, Zambia.

[5]The culture-based regional grouping scheme adapted from the general scheme used by Gupta and colleagues (GLOBE, 2002) Reference includes the following grouping (in CAPITALS) and countries: ANGLO: Australia, Canada, Ireland, New Zealand, South Africa, United Kingdom; ARAB: Egypt, Kuwait, Morocco, Qatar, Turkey; CONFUCIAN ASIA: China, Hong Kong, Japan, Singapore, South Korea; EASTERN EUROPE: Albania, Greece, Hungary, Poland, Russian Federation, Slovenia; EUROPEAN NORDIC: Denmark, Finland, Sweden; GERMANIC: Austria, Germany, Netherlands; LATIN AMERICA: Argentina, Brazil, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Mexico, Venezuela; LATIN EUROPE: France, Israel, Italy, Portugal, Spain, Switzerland; SOUTHERN ASIA: India, Indonesia, Iran, Malaysia, Philippines, Thailand; SUB-SAHARA AFRICA: Namibia, Nigeria, Zambia; COUNTRIES NOT INCLUDED: Algeria, Azerbaijan, Bahrain, Belgium, Benin, Bulgaria, Burkina Faso, Burundi, Chile, Croatia, Czech Republic, Democratic Republic of Congo, Dominican Republic, Estonia, Ethiopia, Fiji, Gambia, Ghana, Guinea, Iraq, Ivory Coast, Jamaica, Jordan, Kenya, Latvia, Lebanon, Liberia, Libya, Lithuania, Luxembourg, Malawi, Mali, Niger, Norway, Oman, Pakistan, Panama, Peru, Romania, Saudi Arabia, Senegal, Serbia and Montenegro, Sierra Leone, Sri Lanka, Sudan, Suriname, Syria, Tanzania, Tunisia, Uganda, United Arab Emirates, Uruguay, Yemen.

[6]The trade and investment-based regional grouping scheme adapted from the general scheme used by Donnenfeld (2003) includes the following groupings (in CAPITALS) and countries: ANDEAN: Chile, Colombia, Ecuador, Peru; ASEAN: Indonesia, Malaysia, Philippines, Singapore, Thailand; EU: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden; MERCOSUR: Argentina, Brazil, Uruguay; NAFTA: Canada, Mexico; PAN-ARAB: Bahrain, Egypt, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Qatar, Saudi Arabia, Sudan, Syria, Tunisia, United Arab Emirates, Yemen; NOT AFFILIATED TO A TRADE AGREEEMENT IN 2000: Albania, Algeria, Australia, Azerbaijan, Benin, Bulgaria, Burkina Faso, Burundi, China, Costa Rica, Croatia, Czech Republic, Democratic Republic of Congo, Dominican Republic, El Salvador, Estonia, Ethiopia, Fiji, Gambia, Ghana, Guatemala, Guinea, Hong Kong, Hungary, India, Iran, Israel, Ivory Coast, Jamaica, Japan, Kenya, Latvia, Liberia, Lithuania, Malawi, Mali, Namibia, New Zealand, Niger, Nigeria, Norway, Pakistan, Panama, Poland, Romania, Russian Federation, Senegal, Serbia and Montenegro, Sierra Leone, Slovenia, South Africa, South Korea, Sri Lanka, Suriname, Switzerland, Tanzania, Turkey, Uganda, United Kingdom, Venezuela, Zambia.

[7]The trade and investment-based regional grouping scheme adapted from the general scheme used by Rugman and Verbeke (2004) includes the following groupings (in CAPITALS) and countries: ASIA PACIFIC:

Australia, Azerbaijan, Bahrain, China, Hong Kong, India, Indonesia, Iran, Iraq, Israel, Japan, Jordan, Kuwait, Lebanon, Malaysia, New Zealand, Oman, Pakistan, Philippines, Qatar, Saudi Arabia, Singapore, South Korea, Sri Lanka, Syria, Thailand, Turkey, United Arab Emirates, Yemen; EUROPE: Albania, Austria, Belgium, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Norway, Poland, Portugal, Romania, Russian Federation, Serbia and Montenegro, Slovenia, Spain, Sweden, Switzerland, United Kingdom; NORTH AMERICA: Canada, Mexico; COUNTRIES NOT INCLUDED IN EXTENDED TRIAD: Algeria, Argentina, Benin, Brazil, Burkina Faso, Burundi, Chile, Colombia, Democratic Republic of Congo, Costa Rica, Dominican Republic, Ecuador, Egypt, El Salvador, Ethiopia, Fiji, Gambia, Ghana, Guatemala, Guinea, Ivory Coast, Jamaica, Kenya, Liberia, Libya, Malawi, Mali, Morocco, Namibia, Niger, Nigeria, Panama, Peru, Senegal, Sierra Leone, South Africa, Sudan, Suriname, Tanzania, Tunisia, Uganda, Uruguay, Venezuela, Zambia.

[8]Both measures require group redefinition as a polygon closely approximating the actual shape. To measure contiguity, we then compare all pairs of perimeter points on a group polygon and choose the pair yielding the longest geodesic distance. To measure compactness, we use the following mathematical expression: Compactness=$4\pi a/p^2$, where a is the area and p is the perimeter of the group polygon.

[9]Mathematically, MF-A Ps R^2 can be summarized in the following expression:

$$\text{Pseudo } R^2_{\text{adj}} = 1 - \frac{\ln \hat{L}(M_{\text{full}}) - k}{\ln \hat{L}(M_{\text{intercept}})}$$

where $\ln \hat{L}(M_{\text{full}})$ is the estimated log likelihood for a model with all parameters, k, and $\ln \hat{L}(M_{\text{intercept}})$ is the estimated log likelihood for the same model with an intercept only.

[10]Mathematically, AIC can be summarized in the following expression:

$$\text{AIC} = 2k - 2\ln\left(\hat{L}\right)$$

where k is the number of parameters, and $\ln\left(\hat{L}\right)$ is the estimated log likelihood. Let $\text{AIC}_{\text{minimum}}$ be the "best" (minimizes information loss) model in a class of models; then the likelihood that another model i in the same class also minimizes information loss is given by $\exp[(\text{AIC}_{\text{minimum}} - \text{AIC}_i)/2]$.

[11]ESS is calculated according to the following expression:

$$ESS = \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2$$

where n is the number of observations, and x_i is the value of observation i.

[12]Mathematically, the Metropolis criterion for accepting non-intuitive changes in a partition, $P_{s\ new}$, can be summarized in the following expression: $P_{Accept} = \min(1, e^{-\Delta ESS/T})$. P_{Accept} is the 0–1 probability that the simulated annealing algorithm will replace old partition, $P_{s\ old}$, with new partition, $P_{s\ new}$. The acceptance (replacement) probability is 1 (certain) if $ESS_{new} < ESS_{old}$. If $ESS_{new} > ESS_{old}$, then acceptance depends on comparison of a pseudo-randomly generated value R, where $0 < R < 1$ to number to $e^{-\Delta ESS/T}$. ΔESS is the difference between ESS_{old} and ESS_{new}. T is the temperature of the annealed system. If $e^{-\Delta ESS/T} > R$ then a non-intuitive ($ESS_{new} > ESS_{old}$) move is accepted, and the simulated annealing algorithm moves "uphill". Lower T and/or larger ΔESS renders this inequality less likely, and thus acceptance of non-intuitive moves less likely.

[13]We illustrate this process in a figure available electronically at: http://www.csom.umn.edu/faculty-research/vaal0001/Paul_M_Vaaler.aspx

[14]The Mathematica and Matlab program codes are available electronically at: http://www.csom.umn.edu/faculty-research/vaal0001/Paul_M_Vaaler.aspx

[15]We illustrate regional grouping scheme refinement processes based on simulated annealing in maps for each of seven schemes evaluated. They are available electronically at: http://www.csom.umn.edu/faculty-research/vaal0001/Paul_M_Vaaler.aspx

[16]We obtain similar results supporting Hypothesis 1 when we replace MF-A Ps R^2 with an alternative pseudo-R^2 proposed by Estrella (1995), and when we replace AIC with an alternative Bayesian information criterion proposed by Schwarz (1978). These results are available electronically at: http://www.csom.umn.edu/faculty-research/vaal0001/Paul_M_Vaaler.aspx

REFERENCES

Aerts, J. C., & Heuvelink, G. B. 2002. Using simulated annealing for resource allocation. *International Journal of Geographical Information Science*, 16(6): 571–587.

Agnew, J. 1999. Regions on the mind does not equal regions of the mind. *Progress in Human Geography*, 23(1): 91–96.

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6): 716–723.

Alrefaei, M. H., & Andradóttir, S. 1999. A simulated annealing algorithm with constant temperature for discrete stochastic optimization. *Management Science*, 45(5): 784–764.

Angel, J. L. 2001. *Directory of American firms operating in foreign countries*. New York: Simon & Schuster.
Antràs, P., Desai, M., & Foley, F. 2007. Multinational firms, FDI flows and imperfect capital markets. *Quarterly Journal of Economics*, 124(3): 1171–1219.
Arregle, J. L., Miller, T. L., Hitt, M. A., & Beamish, P. W. 2013. Do regions matter? An integrated institutional and semiglobalization perspective on the internationalization of MNEs. *Strategic Management Journal*, first published online 22 March. doi: 10.1002/smj.2051.
Asmussen, C. G. 2009. Local, regional or global? Quantifying MNE geographic scope. *Journal of International Business Studies*, 40(7): 1192–1205.
Banalieva, E. R., & Dhanaraj, C. 2013. Home-region orientation in international expansion strategies. *Journal of International Business Studies*, 44(2): 89–116.
Banalieva, E. R., & Eddleston, K. A. 2011. Home-region focus and performance of family firms: The role of family vs non-family leaders. *Journal of International Business Studies*, 42(8): 1060–1072.
Bartlett, C. A., & Ghoshal, S. 1991. Global strategic management: Impact on the new frontiers of strategy research. *Strategic Management Journal*, 12(S1): 5–16.
Berkowitz, D., Pistor, K., & Richard, J.-F. 2003. Economic development, legality and the transplant effect. *European Economic Review*, 47(1): 165–195.
Beugelsdijk, S., McCann, P., & Mudambi, R. 2010. Introduction: Place, space and organization – Economic geography and the multinational enterprise. *Journal of Economic Geography*, 10(4): 485–493.
Breschi, S., & Lissoni, F. 2009. Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows. *Journal of Economic Geography*, 9(4): 439–468.
Buckley, P. J., & Ghauri, P. N. 2004. Globalisation, economic geography and the strategy of multinational enterprises. *Journal of International Business Studies*, 35(2): 81–98.
Carley, K. M., & Svoboda, D. M. 1996. Modeling organizational adaptation as a simulated annealing process. *Sociological Methods & Research*, 25(1): 138–168.
Clark, T., & Knowles, L. L. 2003. Global myopia: Globalization theory in international business. *Journal of International Management*, 9(4): 361–372.
Clark, T., Knowles, L. L., & Hodis, M. 2004. Global dialogue: A response to the responders in the special globalization issue of JIM. *Journal of International Management*, 10(4): 511–514.
Clougherty, J., & Grajek, M. 2008. The impact of ISO 9000 diffusion on trade and FDI: A new institutional analysis. *Journal of International Business Studies*, 39(4): 613–633.
Cuervo-Cazurra, A. 2008. The effectiveness of laws against bribery abroad. *Journal of International Business Studies*, 39(4): 634–651.
DCE. 2012. *Dictionary of contemporary English*. Longman, http://www.ldoceonline.com/dictionary/region. Accessed 15 June 2012.
Donnenfeld, S. 2003. Regional blocs and foreign direct investment. *Review of International Economics*, 11(5): 770–788.
Dunning, J. H. 2000. *Regions, globalization, and the knowledge-based economy*. Oxford: Oxford University Press.
Dyck, A., Volchkova, N., & Zingales, L. 2008. The corporate governance role of the media: Evidence from Russia. *Journal of Finance*, 63(2): 1093–1135.
Estrella, A. 1995. *Measures of fit with dichotomous dependent variables: Critical review and a new proposal*. New York: Federal Reserve Bank of New York.
Fawn, R. 2009. "Regions" and their study: Wherefrom, what for and whereto? *Review of International Studies*, 33(S1): 5–34.
Flores, R. G., & Aguilera, R. V. 2007. Globalization and location choice: An analysis of US multinational firms in 1980 and 2000. *Journal of International Business Studies*, 38(7): 1187–1210.
Fortune. 2001. The fortune 500: Our annual list of the largest US corporations. *Fortune*, 143(8): F1–F23.
Fox, I., Srinivasan, S., & Vaaler, P. 1997. A descriptive alternative zto cluster analysis: Understanding strategic group performance with simulated annealing. In M. Ghertman, J. Obadia, & J.L. Arregle (Eds) *Statistical models for strategic management*. Norwell, MA: Kluwer Academic Publisher.
Fratianni, M. 2009. The gravity equation in international trade. In A.M. Rugman (Ed) *The Oxford handbook of international business*. Oxford: Oxford University Press.
Fratianni, M., & Oh, C. H. 2009. Expanding RTAs, trade flows and the multinational enterprise. *Journal of International Business Studies*, 40(7): 1206–1277.
Friedman, T. L. 2005. *The world is flat: A brief history of the twenty-first century*. New York: Farrar, Straus and Giroux.
Fryer, R. G., & Holden, R. 2011. Measuring the compactness of political districting plans. *Journal of Law & Economics*, 54(3): 493–535.
Fukuyama, F. 1992. *The end of history and the last man*. New York: Avon Books.
Furnham, A., Kirkcaldy, B., & Lynn, R. 1994. National attitudes to competitiveness, money and work amongst young people: First, second and third world differences. *Human Relations*, 47(1): 119–132.
Gatignon, H., & Kimberly, J. R. 2004. Globalization and its challenges. In H. Gatignon, J.R. Kimberly, & R.E. Gunther (Eds) *The INSEAD-Wharton alliance on globalizing strategies for building successful global business*. Cambridge: Cambridge University Press.
Geringer, J. M., Beamish, P. W., & da Costa, R. C. 1989. Diversification strategy and internationalization: Implications for MNE performance. *Strategic Management Journal*, 10(2): 109–119.
Ghemawat, P. 2003. Semiglobalization and international business strategy. *Journal of International Business Studies*, 34(2): 138–152.
Ghoshal, S. 1987. Global strategy: An organizing framework. *Strategic Management Journal*, 8(5): 425–440.
Goffe, W. L., Ferrier, G. D., & Rogers, J. 1994. Global optimization of statistical functions with simulated annealing. *Journal of Econometrics*, 60(1–2): 65–99.
Gupta, V., Hanges, P. J., & Dorfman, P. 2002. Cultural clusters: Methodology and findings. *Journal of World Business*, 37(1): 11–15.
Han, B. 1994. Optimal file management for a stage system using magnetic and optical disks. *Information and Decision Technologies*, 19(3): 393–412.
Hofstede, G. 2001. *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*, (2nd edn). Thousand Oaks, CA: Sage.
House, R., Javidan, M., Hanges, P. J., & Dorfman, P. 2002. Understanding cultures and implicit leadership theories across the globe: An introduction to project GLOBE. *Journal of World Business*, 37(1): 3–10.
Katrishen, F. A., & Scordis, N. A. 1998. Economies of scale in services: A study of multinational insurers. *Journal of International Business Studies*, 29(2): 305–323.
Kirkman, B. L., Lowe, K. B., & Gibson, C. B. 2006. A quarter century of *Culture's Consequences*: A review of empirical research incorporating Hofstede's cultural values framework. *Journal of International Business Studies*, 37(3): 285–320.
Kirkpatrick, S., Gellatt, C., & Vecchi, M. 1983. Optimization by simulated annealing. *Science*, 220(4598): 671–680.
Kogut, B. 2009. Methodological contributions in international business and the direction of academic research activity. In A.M. Rugman (Ed) *The Oxford handbook of international business*. Oxford: Oxford University Press.
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.
Kwok, C. C.-Y., & Tadesse, S. 2006. National culture and financial systems. *Journal of International Business Studies*, 37(2): 227–247.
La Porta, R., López-de-Silanes, F., & Shleifer, A. 2008. The economic consequences of legal origins. *Journal of Economic Perspectives*, 46(2): 285–332.
Lewis, M. W., & Wigen, K. 1997. *The myth of continents: A critique of metageography*. Berkeley, CA: University of California Press.

Long, S. J., & Freese, J. 2006. *Regression models for categorical dependent variables using Stata*. College Station, TX: Stata Press.
Madhavan, R., & Iriyama, A. 2009. Understanding global flows of venture capital: Human networks as the "carrier wave" of globalization. *Journal of International Business Studies*, 40(8): 1241–1259.
McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed) *Frontiers in econometrics*. New York: Academic Press.
Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., & Teller, E. 1953. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6): 1087–1092.
Nachum, L., Zaheer, S., & Gross, S. 2008. Does it matter where countries are? Proximity to knowledge, markets and resources, and MNE location choices. *Management Science*, 54(7): 1252–1265.
Niemi, R., Groffman, B., Calucci, C., & Hofeller, T. 1990. Measuring compactness and the role of a compactness standard in a test for partisan and racial gerrymandering. *The Journal of Politics*, 52(4): 1155–1181.
North, D. C. 1990. *Institutions, institutional change and economic performance*. New York: Cambridge University Press.
Ohmae, K. 1985. *Triad power: The coming sharp of global competition*. New York: Free Press.
Osegowitsch, T., & Sammartino, A. 2008. Reassessing (home-) regionalisation. *Journal of International Business Studies*, 39(1): 184–196.
Poon, J. 1997. The cosmopolitanization of trade regions: Global trends and implications, 1965–1990. *Economic Geography*, 73(4): 390–404.
Reynolds, T. H., & Flores, A. 1989. *Foreign law: Current sources of codes and basic legislation in jurisdictions in the world*. Littleton, CO: F. B. Rothman.
Ronen, S., & Shenkar, O. 1985. Clustering countries on attitudinal dimensions: A review and synthesis. *Academy of Management Review*, 10(3): 435–454.
Rugman, A. M., & Oh, C. H. 2012. Why the home region matters: Location and regional multinationals. *British Journal of Management*, first published online 28 February, doi: 10.1111/j.1467-8551.2012.00817.x.
Rugman, A. M., & Verbeke, A. 2004. A perspective on regional and global strategies of multinational enterprises. *Journal of International Business Studies*, 35(1): 3–18.
Rugman, A. M., Li, J., & Oh, C. H. 2009. Are supply chains global or regional? *International Marketing Review*, 26(4/5): 384–395.
Schwarz, G. E. 1978. Estimating the dimension of a model. *Annals of Statistics*, 6(2): 461–464.
Semmler, W., & Gong, G. 1996. Estimating parameters of real business cycle models. *Journal of Economic Behavior & Organization*, 30(3): 301–325.
Simmons, B. A., Elkins, Z., & Guzman, A. 2006. Competing for capital: The diffusion of bilateral investment treaties, 1960–2000. *International Organization*, 60(4): 811–846.
Singh, J. 2005. Collaborative networks as determinants of knowledge diffusion patterns. *Management Science*, 51(5): 56–70.
Slangen, A., & Beugelsdijk, S. 2010. The impact of institutional hazards on foreign multinational activity: A contingency perspective. *Journal of International Business Studies*, 41(6): 980–995.
Slangen, A., Beugelsdijk, S., & Hennart, J.-F. 2011. The impact of cultural distance on bilateral arm's length exports: An international business perspective. *Management International Review*, 51(6): 875–896.
Sorenson, O., & Baum, J. A. C. (eds) 2003. Editors' introduction: Geography and strategy: The strategic management of space and place. In *Geography and strategy*. Advances in Strategic Management. Vol. 20: 1–19. Bingley, UK: Emerald Group.
The MathWorks. 2012. *MATLAB 7.14 (R2012a)*. Natick, MA: The MathWorks Inc.
United Nations (UN). 2007. *Composition of macro geographical (continental) regions, geographical sub-regions, and selected economic and other groupings*. http://unstats.un.org/unsd/methods/m49/m49regin.htm, accessed 30 July 2012.
UNCTAD. 2002. *World investment report 2002: Transnational corporations and export competitiveness*. New York: United Nations.
Vaaler, P. M. 2008. How do MNCs vote in developing country elections? *Academy of Management Journal*, 51(1): 21–43.
Vaaler, P. M. 2011. Immigrant remittances and the venture investment environment of developing countries. *Journal of International Business Studies*, 42(9): 1121–1149.
Vaaler, P. M., & McNamara, G. 2004. Crisis and competition in expert organizational decision-making: Credit rating agencies and their response to turbulence in emerging economies. *Organization Science*, 15(6): 687–703.
Wagenmakers, E. J., & Farrell, S. 2004. AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*, 11(1): 192–196.
Walter, B., & Bernard, F. 1978. Ash pile or rising phoenix: A review of the status of regional geography. *Journal of Geography*, 77(5): 192–197.
Wei, S. J. 2000. How taxing is corruption on international investors? *The Review of Economic and Statistics*, 82(1): 1–11.
Wolfram Research. 2012. *Mathematica edition: Version 9.01*. Champaign, IL: Wolfram Research, Inc.
Xu, D., & Shenkar, O. 2002. Institutional distance and the multinational enterprise. *Academy of Management Review*, 27(4): 608–618.
Yip, G. S. 1992. *Total global strategy: Managing for worldwide competitive advantage*. Englewood Cliffs, NJ: Prentice Hall.
Zwinkels, R. C. J., & Beugelsdijk, S. 2010. Gravity equations: Workhorse or Trojan horse in explaining trade and FDI patterns across time and space? *International Business Review*, 19(1): 102–115.

ABOUT THE AUTHORS

Ricardo Flores earned his PhD at the University of Illinois and is currently a Lecturer (Assistant Professor) within the Australian School of Business at the University of New South Wales (Sydney, Australia). His research explores different aspects of the location choices of MNEs. He also does research on international processes of technological evolution. He was born in and is a citizen of Argentina.

Ruth V Aguilera is an Associate Professor at the University of Illinois and holds a joint appointment at the ESADE Business School. She received her PhD in sociology from Harvard University. Her research interests fall at the intersection of economic sociology and international business, specifically in the fields of global strategy and comparative corporate governance. She was born in and is a citizen of Spain.

Paul M Vaaler earned his PhD and is currently an Associate Professor of International Business at the University of Minnesota's Carlson School of Management. His research examines risk and investment behavior by firms and individuals active in emerging-market countries experiencing economic and political modernization. He was born in

and is a citizen of the USA. His email address is vaal0001@umn.edu.

Arash Mahdian did his graduate studies in econometrics at the University of Illinois at Urbana-Champaign, and is now Lead Economics Developer at Wolfram Research, the maker of Mathematica software. His research interests cover various aspects of computational economics, including optimization of complex systems. He was born in Iran and is an Iranian citizen. His email address is arashm@wolfram.com.

Accepted by Sjoerd Beugelsdijk and Ram Mudambi, Guest Editors, 6 March 2013. This paper has been with the authors for three revisions.