HARVARD | BUSINESS | SCHOOL

Shifting Centers of Gravity: Host Country versus Headquarters Influences on MNC Subsidiary Knowledge Inheritance

Prithwiraj Choudhury
Tarun Khanna
Mike Horia Teodorescu

Working Paper 19-020

Shifting Centers of Gravity: Host Country versus Headquarters Influences on MNC Subsidiary Knowledge Inheritance

Prithwiraj Choudhury
Harvard Business School

Mike Horia Teodorescu
Boston College

Tarun Khanna
Harvard Business School

Working Paper 19-020

Shifting Centers of Gravity:
Host Country versus Headquarters Influences on MNC Subsidiary Knowledge Inheritance [1]

Prithwiraj Choudhury [a], Mike Horia Teodorescu [b] and Tarun Khanna [a]

[a] Harvard Business School, Harvard University, Boston MA USA

[b] Carroll School of Management, Boston College, Chestnut Hill MA USA

ABSTRACT

While there is a rich literature of knowledge inflows into the multinational subsidiary, the literature is rooted in how subsidiaries inherit knowledge from the headquarters (HQ). In this paper we take the first step to liberate the construct of "subsidiary knowledge inheritance" from its umbilical attachment to the MNC headquarters. We build on the prior theory of subsidiary absorptive capacity and argue that larger subsidiaries, characterized by greater knowledge stock and a greater fraction of local employees, could plausibly absorb more knowledge from the local host country context compared to absorbing knowledge from the headquarters. We test our theoretical prediction using a novel methodology based on the classic gravity equation in economics and measures of knowledge distance. Using a custom dataset of patents filed by all global subsidiaries of the top 25 patenting U.S. multinationals, we find that the relative influence of the HQ and the host country on knowledge inflows to the subsidiary depends on subsidiary knowledge stock. Our findings show that as knowledge stock at the subsidiary grows, it absorbs more knowledge from the host country compared to the headquarters. We additionally provide a stylized case study of CISCO opening a "second headquarters" in India to exemplify our core proposition.

KEYWORDS: Multinational firms, MNCs, Subsidiary Knowledge Inheritance, Subsidiary Absorptive capacity, Host Context, Gravity Model, Cosine Similarity, Fidelity Similarity, Second Headquarters

[1] Corresponding author – Prithwiraj Choudhury; email: pchoudhury@hbs.edu. Acknowledgments: This work was produced with research funding from Harvard Business School. The authors thank researcher Patrick Clapp from Harvard Business School Baker Research for data processing assistance and valuable feedback on work in progress and Researcher Xiang Ao from Harvard Business School Research Computing Services for early discussions on methodology and valuable feedback. We also thank Juan Alcacer, Daniel Brown, Yo-Jud Cheng, Do Yoon Kim, William Kerr, Jan Rivkin, Jordan Siegel, Haris Tabakovic, Michael Toffel, Dennis Yao, Shane Greenstein, Aaron Yoon, the seminar participants at Halmstad University's CIEL innovation seminar, and the participants at the SMS Berlin Conference for valuable feedback on earlier versions of this work. The second author is also grateful to the HBS Doctoral Seminar for valuable input and to Harvard Business School for the doctoral fellowship during which part of this paper was developed. We also thank Andrew Marder from Harvard Business School Research Computing for feedback.

Scholars have long hypothesized that multinational firms (MNCs) exist because of their ability to transfer and exploit knowledge more effectively and efficiently in the intrafirm context than would be possible through external market mechanisms. As Gupta and Govindrajan (2000) point out, the internalization of the intangible assets argument, originally advanced by Hymer (1960), has been widely accepted as the theory of why MNCs exist.[2] A rich empirical literature studies knowledge "inflows" and "outflows" from the perspective of MNC subsidiaries and examines how MNC headquarters (HQ) and other subsidiaries influence such knowledge flows (Feinberg & Gupta, 2004; Gupta & Govindrajan, 2000; Singh, 2008). In fact, over the past several decades, this literature has viewed the multinational subsidiary as being umbilically attached to the headquarters and to the broader MNC network for its endowment of knowledge. Berry (2015) stresses the importance of "knowledge inheritance" from the MNC headquarters to the subsidiary. As Monteiro, Arvidsson and Birkinshaw (2008) argue, knowledge transfers in MNCs typically occurs between "highly capable" members of the MNC network (the network comprising the headquarters and all subsidiaries). Given this, the "isolated" subsidiary, cutoff from knowledge transfers from the headquarters and other important subsidiaries might suffer the "liability of internal isolation".

In this paper we take the first step towards liberating the construct of "subsidiary knowledge inheritance" from its umbilical attachment to the MNC headquarters. In doing so, we study subsidiary knowledge inheritance from a second source: *the host country context*. The intellectual tradition of studying the importance of host country contexts to MNC subsidiaries dates back to the work of Prahalad, Doz and Bartlett in the 1970s and the 1980s. This stream of research focused on the fact that subsidiaries face two sets of institutional, environmental and relational forces, one each from the headquarters and the local context, and this realization led to insights around the importance of

[2] Buckley and Casson (1976); Caves (1971), Caves, Christensen, and Diewert (1982); Ghoshal (1987); Ghoshal and Bartlett (1990); Kindleberger (1969); Porter (1986); Teece (1981).

"national differentiation" (Bartlett 1986) and "pressures for local responsiveness" (Prahalad and Doz, 1987).[3] Yet two important facts need to be kept in mind: firstly, even when scholars have examined the importance of the local context to the subsidiary, it has mostly been in relation to how the local context helps or hurts knowledge transfers from the headquarters to the subsidiary. As an example, while building on the institutional and relational duality faced by the subsidiary, Kostova and Roth (2002) study how the transfer of knowledge (of organizational practices) from the headquarters to the subsidiary is related to the host country's regulatory, cognitive and normative institutional profiles. In other words, prior scholarship has not theorized about whether and why knowledge inheritance to the subsidiary *from* the host country could be comparable or even exceed knowledge inheritance from the headquarters. The second fact to keep in mind is that even important empirical studies on knowledge flows within MNCs, such as Gupta and Govindrajan (2000) and Feinberg and Gupta (2004), focus on knowledge flows between the subsidiary and the HQ or other subsidiaries and ignore knowledge flows from the host country to the subsidiary. This gap exists even though recent literature in strategy and international business has made a strong case for studying how the local context shapes the multinational subsidiary (Khanna, Palepu, & Sinha 2005; Khanna & Palepu, 2010; Mudambi, Pedersen,& Andersson, 2014; Meyer, Mudambi, & Narula, 2014; Santos & Williamson, 2015). Within this nascent literature, Meyer and Estrin (2013) argue that host country resources such as local knowledge and local talent are critical for MNC subsidiaries. This stream of research also argues that these host country resources evolve over time, especially in rapidly growing emerging markets. However, an important white space in the literature has been the lack of scholarship comparing subsidiary knowledge inheritance from the two important knowledge pools available to the subsidiary, i.e. the headquarters and the host country context. We also lack insights on what subsidiary level

[3] There is a long subsequent tradition of studying the importance of local contexts to subsidiaries; scholars in this tradition have posited that MNC subsidiaries have to adapt to the local context (Birkinshaw & Hood, 1998; Ghoshal & Nohria, 1989; Kostova & Roth, 2002; Nobel & Birkinshaw, 1998; Nohria & Ghoshal, 1994; Rugman & Verbeke, 2001).

factors might tilt the relative balance of subsidiary knowledge inheritance from predominantly inheriting knowledge from the headquarters to predominantly inheriting knowledge from the local context. *This is the theoretical and empirical gap we seek to fill.*

To do so, we build on prior theory of subsidiary absorptive capacity and posit that *larger subsidiaries, characterized by greater knowledge stock, are likely to absorb more knowledge from the local host country context compared to absorbing knowledge from the headquarters*. Underlying this theoretical proposition are two core assumptions: firstly we build on the literature of subsidiary talent distribution literature (Chang, Gong and Peng, 2012; Choudhury 2015) to assume that subsidiaries that accumulate a greater knowledge stock are likely to be characterized by a talent pool where local employees are in the majority, compared to expatriates deputed from the headquarters. Secondly, we build on Cohen and Levinthal (1990) to argue that the absorptive capacity of local employees working at the subsidiary is likely to be aligned with knowledge where they have prior familiarity and as an extension, local employees are likely to draw upon pools of knowledge available within the context of the host country.

An important limitation related to studying our core theoretical proposition is the lack of an empirical technique capable of conducting an "apples to apples" comparison of how the HQ and the host country influence knowledge flows to the subsidiary. Such an empirical comparison would need to be conducted on two comparable empirical "arms." The first arm would estimate the influence of the headquarters on knowledge flows to the subsidiary, while the second would measure, using the same theoretical framework, the influence of the host country context on knowledge flows to the subsidiary. To work around this difficulty, we use the gravity model in economics to estimate two comparable specifications – one that measures knowledge flows from the headquarters to a multinational subsidiary and another that measures knowledge flows from the host country context to the subsidiary. The gravity model in economics was introduced in the trade literature with Tinbergen's (1962) work and was later formalized by Anderson (1979). Intuitively, this gravity model builds on the original model of gravity in Newtonian mechanics, where the force of attraction between two bodies is proportional to the masses of the two bodies and inversely proportional to the square of the distance between those bodies. The equivalent form of the gravity equation in the trade literature substitutes the force of attraction as a dependent variable with a measure of trade, such as the dollar flow of traded goods between two countries (Anderson, 1979) in relation to the masses of the two countries as measured in GDP (Mátyás, 1997), and in relation to a distance variable, often specified as the geographic distance between the two trading regions (Anderson, 1979).[4] The gravity model is highly generalizable, as long as one has sensible measures of flow (force), mass, and distance. Consequently, it is now being used in various fields beyond trade economics.[5] For example, Lewer and Van den Berg (2008) developed a gravity model of immigration. In this case, the flow is immigration, the mass variables are the populations of the pairs of countries, and the distance is defined traditionally as geographic distance.

We apply a generalized form of the gravity model to study the relative influence of the MNC headquarters and the local context of the host country on knowledge inflows to the MNC subsidiary. Given that our study focuses on "subsidiary knowledge inheritance", we focus on knowledge inflows into the subsidiary. To conduct this empirical analysis, we estimate two "gravity" equations. The first equation uses the "mass" of the headquarters, the "mass" of the subsidiary and the "distance" between

[4] Bergstrand (1985) further expanded the theoretical foundations of the gravity model in the trade literature by deriving it from a general equilibrium model. While gravity equation-based regression models have generally been estimated via OLS (e.g., Mátyás, 1997), the OLS estimator has been shown to be biased in the case of heteroskedasticity (Silva & Tenreyro, 2006). Silva and Tenreyro (2006) instead proposed a Poisson pseudo-maximum-likelihood estimator for the cases where the errors are heteroskedastic. Additional econometric developments have been made, for example by Baier and Bergstrand (2008).

[5] Other relevant papers that use the gravity model include Anderson and Wincoop (2003), who focus on the effect of national borders as trade barriers and determine that borders substantially reduce trade by 20-50%. Waugh (2010) focuses on determining trade flow asymmetries between countries based on differences in the standards of living of the trading countries. Summary (1989) used measures of political factors between trading partners as additional independent variables to augment a regression model based on the gravity equation.

them; the second equation uses the "mass" of the host country, the "mass" of the subsidiary and the "distance" between them.

To estimate these two equations, we need relevant measures of knowledge flows from the MNC headquarters and the host country to the MNC subsidiary, measures for the "mass" of the knowledge stock at the MNC headquarters/in the local context and measures for the "knowledge distance" between the focal subsidiary and the MNC headquarters/local context. We quantify knowledge flows through the use of patent citations, a widely accepted measure in the innovation literature (Jaffe *et al.*, 1993). To estimate the "mass" variables, we use the stock of patents filed by an entity. Measuring "distance" between MNC entities (HQ, subsidiary and host country context) presents unique challenges: the gravity model in economics has typically used geographic distance to model trade flows. However, the literature in knowledge flows has shown that social and ethnic ties between inventors influence knowledge flows between locations (Agrawal *et al.*, 2006; Agrawal *et al.*, 2008). Given this, geographic distance may be less salient in determining knowledge flow between MNC entities. Instead, we use several novel measures of distance based on cosine similarity measures. As we explain in detail later, we also employ a novel measure of distance based on the Bhattacharya coefficient, and "fidelity similarity" (Deza & Deza, 2015). Using these measures, we separately measure the effect of the MNC headquarters and the local context on knowledge inflows to the subsidiary.

To conduct this analysis, we created and used a novel dataset of U.S. patents filed by *all* subsidiaries of the top 25 US headquartered multinationals at the US Patent and Trademark Office (USPTO) over a period of 7 years (2005–2011). These patents were filed at either the headquarters or any of the subsidiaries of these firms around the world. We coded the location from the location of the inventors who were listed on each patent. We also created a "mass" measure for each headquarter, host country and subsidiary from 2005 to 2011 and created distance measures using custom-written software code. We report several results. First, we estimate the gravity model specifications separately for the influence of the headquarters on knowledge flows into the subsidiary and the influence of the host country on knowledge flows into the subsidiary: as expected the mass and distance measures are correlated to knowledge flows. These results are robust to the inclusion of several controls. Second and crucial for our paper, we conduct analyses to establish that the relative influence of the HQ and the host country on subsidiary knowledge inflows depends on the size of the subsidiary. As the size of the subsidiary grows (measured using the stock of patents held by the subsidiary), the relative influence of the host country on subsidiary knowledge inheritance increases compared to knowledge inheritance from the headquarters. This indicates possible heterogeneity across MNC subsidiaries in the relative importance of the host country context and the headquarters. Our results indicate that from the perspective of subsidiary knowledge inheritance, the host country context is more important than the role of the headquarters for subsidiaries that have a greater stock of patented innovations.

Our results contribute to several streams of the multinational literature including research on multinational subsidiary knowledge inheritance, changes in the subsidiary mandate/charter, and the importance of the host country context. In fact, our research introduces the gravity model to the strategy literature and makes a theoretical contribution by identifying conditions related to a *shifting center of gravity* of subsidiary knowledge inheritance: from the headquarters to the host country context. Our research also makes an empirical contribution, by introducing novel measures of "knowledge distance" to the strategy and IB literature. For managers of multinational firms, our research indicates that in the current environment of skepticism towards the role of multinationals around the world and the perceived trends towards nationalism and "de-globalization", being tethered to the host country can provide subsidiaries with a lifeline of knowledge inheritance.

THEORETICAL FOUNDATIONS AND HYPOTHESES

Theoretical Foundations – The Headquarters

The earliest models of the MNC view the organizational structure as a "centralized hub" (Bartlett, 1986), where the HQ directs resources, tasks, and relationships to the MNC subsidiaries. Ghoshal and Bartlett (1990) introduced the "interorganizational network" view of the MNC, where the organizational units of the MNC, which include its subsidiaries and its HQ, are embedded in a "network" that consists of all the entities the MNC interacts with. Within this interorganizational network, the HQ may assign different strategic roles to its MNC subsidiaries. There is a rich subsequent literature establishing the pivotal, even perhaps *paternalistic* role of the HQ in the MNC organization (Andersson, Forsgren, and Holm, 2002; Björkman, Barner-Rasmussen, and Li, 2004; Dacin, Beal, and Ventresca, 1999; Ghoshal and Bartlett, 1990; Ghoshal and Nohria, 1989; Gupta and Govindarajan, 1991; and Nell and Ambos, 2013). Ghoshal and Bartlett (1990) see the MNC as "somewhere between […] unitary and federative structures" (Ghoshal and Bartlett 1990, p. 607), meaning that in some MNCs, the goals are set and the decisions are made with full authority by the HQ, with the subsidiaries following directions from the headquarters; while in other MNCs, the subsidiaries are given the choice of whether to ratify the decisions handed to them from the HQ.

The existential importance of the headquarters to the MNC subsidiary permeated the literature on knowledge flows to the MNC. A detailed analysis in Gupta and Govindarajan (2000) shows that "knowledge inflows into a subsidiary is positively associated with richness of transmission channels and motivational disposition to acquire knowledge" (Gupta and Govindajaran, 2000, p. 473); these are both influenced by the HQ, which may invest in the increase of the bandwidth of the transmission channels and force goals for the subsidiary that would motivate it to absorb knowledge. As Gupta and Govindarajan (1991) and Narula (2014) show, different subsidiaries may have very different types of inward knowledge flows and different levels of control exerted by the MNC, which, as argued first in

Gupta and Govindarajan (1991), may be due to the different contexts where the subsidiaries operate. Among the control means that the HQ may use to increase knowledge transfer to itself, Björkman *et al.* (2004) first identify "the specification of knowledge transfer as a criterion of subsidiary performance" (p. 446), which may push the subsidiary to transfer more knowledge to the HQ and other subsidiaries – and to receive more knowledge from other subsidiaries. Björkman *et al.* (2004), Gupta and Govindarajan (2000), and O'Donnell (2000) identify managerial socialization as an opportunity to bring the subsidiary closer to the HQ vision, including through "international training programmes, by establishing international task forces and committees, and by encouraging visits across MNC units" (Björkman *et al.*, 2004, p. 451). Bouquet and Birkinshaw (2008) find that subsidiaries attract the attention of headquarters using their structural position, and by taking initiative and building their profile. They also find that subsidiaries geographically further from their headquarters show stronger initiative, which in turn yields positive attention from the headquarters. In a recent paper, Berry (2015) stresses the importance of knowledge inheritance from the MNC headquarters to the subsidiary. Berry uses unique U.S. Bureau of Economic Analysis data and subsidiary royalty payments to the parent to measure knowledge inflows from the headquarters to the subsidiary and finds that the transfer of knowledge from the headquarters to the subsidiary leads to greater return on asset generation at the subsidiary, especially when there are important clusters of innovation in the home country of the MNC and when the knowledge is transferred to subsidiaries based in host countries that are lagging behind in innovation. In summary, the fact that the HQ influences knowledge flows and knowledge inheritance of the MNC subsidiary is a well-researched topic.

Theoretical Foundations – The Subsidiary

While the initial theorizing in the MNC literature was primarily focused on the central importance of the headquarters, subsequent research viewed the MNC subsidiary as being part of a triadic relationship involving the HQ, the subsidiary and the host country context. This literature also identified various pathways in which the MNC subsidiary could evolve. However, even in this literature development, the HQ arguably did not lose its pre-eminence in either the theoretical or empirical streams of literature.

It has been shown that subsidiaries with higher knowledge output and more connections to their local context are more valuable to the MNC and its *HQ* (Almeida & Phene, 2004). Gupta and Govindarajan (1991), building on the Transaction Cost Economics literature, define this type of subsidiary as playing a "Global Innovator role," in which the subsidiary's benefit to the MNC is driven by its unique knowledge-generating potential, knowledge that is used as currency in exchanges with other units within the MNC organization, as well as by the lower intra-organizational knowledge transfer transaction costs. Monteiro, Arvidsson, and Birkinshaw (2008) have argued that "some subsidiaries are isolated from knowledge transfer activities within the multinational" (p. 90) because they do not belong to the units "perceived to be highly capable" (p. 90), or because of the low "levels of communication and reciprocity" (p. 94). These "levels of communication" are directly connected to the high-capacity channels described by Narula (2014) and to the richness of transmission channels emphasized by Gupta and Govindarajan (2000). Studying the product flow only, Birkinshaw and Morrison (1995) found in a study focused on configurations of the MNCs that the parent-subsidiary relationship differs substantially for "world mandate subsidiaries" and local subsidiaries, with the former experiencing a significantly larger strategic autonomy, which may positively influence the ability of the subsidiary to choose its level of knowledge absorption and knowledge sources (Birkinshaw & Morrison, 1995, p. 744).

In fact, drawing on the theme of different pathways that MNC subsidiaries can undertake in their development and growth, Birkinshaw and Hood (1998) draw attention to the fact that there is an "enormous variety of subsidiaries in existence" (p. 773). The authors employ two theoretical lenses, that of network theory and that of decision processes in large organizations. They conclude that the subsidiaries are continuously evolving as elements of a network, sometimes going beyond a strict dyadic relationship with the HQ, with the evolution being propelled by the "underlying capabilities" of the subsidiaries (Birkinshaw & Hood, 1998, p. 782) and by the degree of autonomy they have to make decisions and take initiative as entities. This tension, between the mandate given to the subsidiary by its parent and the subsidiary's evolving charter, is fueled by the subsidiary's growing capabilities in its *local network*. Subsidiaries are no longer "resource seeking, market seeking, or efficiency seeking" entities (Birkinshaw & Hood, 1998, p. 773); instead, they create their own dynamic capabilities (Nelson & Winter, 1982) and become sources of competitive advantage for their parent organizations. In other words, while the subsidiary charter evolution literature still views the role of the HQ with primacy, there is a gradual realization that the host country context can shape MNC subsidiary charter evolution as well.

Theoretical Foundations – The Host Country

One of the earliest views of the role of host country in the literature is that of the Product Life Cycle model (Vernon, 1966), in which locating a manufacturing plant abroad is a natural part of the product life cycle: later-stage technologies can benefit from the lower cost of manufacturing in a location other than where they were invented. However, the move abroad in this model is driven by production costs, and the subsidiary is always under the directive of the HQ. The later network model of Ghoshal and Bartlett (1990) and Rugman and Verbeke (1992) allows for ties between the subsidiary and the host country that are inherently valuable, external to the HQ and costly for the HQ to develop abroad. Subsequent scholars such as Dunning (1988), Anand and Delios (2002) and Cantwell and Mudambi (2005) have argued that the MNC has to integrate its firm-specific assets (the "ownership" advantages) with location-specific assets of host country contexts (the "location" advantages). Prior research has focused on the importance of factor endowments of the host country, such as the level of technological sophistication of the host country as a factor that influences the decision choice of MNCs to establish R&D centers in certain host countries (Kuemmerle 1999; Feinberg and Gupta, 2004; Alcacer and Zhao, 2012). In fact, Feinberg and Gupta (2004) argue that in addition to host country factor endowments, the ability to capture and utilize knowledge spillovers from competitors located in the host country might drive the R&D location decisions of MNCs.

The most recent studies of knowledge flows and innovation in MNCs build upon this growing role and influence of the host country for the subsidiary, and introduce the perspective that subsidiaries evolve and might turn to their host countries, as opposed to the HQ, to define their charter and role. Meyer, Mudambi and Narula (2011) outline two mechanisms in which the local context can shape an MNC subsidiary – through the quality of its formal and informal institutions (Rodrik et al., 2004) and through resource endowment for MNC subsidiaries.[6] Most important among these resources is knowledge. In fact, the strategy and IB literature has long studied the importance of knowledge inheritance from the host country, in relation to studying why MNCs establish R&D centers in some host countries and not others. Mudambi, Pedersen, and Andersson (2014) find that a subsidiary may turn toward to the national agents of innovation as a replacement of the HQ as a source of ideas and platform for ideas exchange. Mudambi *et al.* (2014) stress that "there is evidence that headquarters' fiat power in MNCs is not absolute" (p. 102). In fact, knowledge assets in the host country may become a basis of the subsidiaries' power. In an analysis of the influence of the MNC and the host country on innovation in the process of knowledge creation by subsidiaries, Almeida and Phene (2004) argue that foreign subsidiaries of the MNC evolve in two contexts: the context of the MNC network and that of the local context. Accordingly, several factors play a role in the generation of innovation by subsidiaries. Two of these factors are identified by Almeida and Phene (2004) at the

[6] Chacar, Newburry and Vissa (2010) show that the quality of formal institutions in the product, labor and financial markets is related to performance persistence of MNCs.

local level, namely local technological richness and the strength of the links the subsidiary has with local entities. O'Donnell (2000) finds that a subsidiary with "a high level of specialized information [the] headquarters does not have" gains a "strategic role" in the MNC (p. 527); such a role provides favorable terms to the management of the subsidiary and is inherently valuable. In summary, subsidiaries have the option to turn toward the local context when the host country provides more resources and more favorable incentives for innovation and knowledge exchange than does the HQ.

The Gravity Model

Based on this exposition of the twin sources of knowledge inheritance for the subsidiary, it is tempting to study the relative influences of the HQ and host country in subsidiary knowledge inheritance. However, the strategy and IB fields lack a theoretical model to conduct such an exploration. To circumvent this gap, we propose the use of the gravitational model as a theoretical and empirical tool to compare the relative influences of the HQ and the host country on knowledge inflows to subsidiaries. Below, we describe the theoretical foundations of the gravity model.

The traditional gravity model in the trade literature establishes an inverse proportionality between a trade variable (such as trade flow between two countries) and the physical distance between the countries, as well as a direct proportionality of the trade variable to two "mass" variables that represent measures of the trading capacity of the two countries, such as their respective GDPs. An example is (Mátyás, 1997, p. 363):

$$\ln(EXP_{ijt}) = \alpha_i + \beta_1 \cdot GDP_{it} + \beta_2 \cdot GDP_{jt} + \beta_3 \cdot d_{ijt} + \epsilon_{ijt}$$

The choice of a distance function and the economic measure equivalent to mass has varied depending on the application. In Mátyás (1997), for example, the dependent variable was the volume of trade between the two countries, and the two masses were the populations of the two countries. In Bergstrand (1985), the distance was the physical distance between economic centers, and the masses were the GDP values in year *t* of the two countries. Without changing the key components of the gravity equation, specifically the two measures of mass and the measure of distance, numerous alterations to the model are possible. This includes the addition of other independent variables such as foreign currency reserves (Mátyás, 1997). This flexibility and the derivation of the regression model encourage wide applications, specifically enabling various definitions of distance (there are many distance functions beyond the popularly used geographic distance) and various options for the independent variable. To shed further light on these variations, one may look at the Newtonian physics formula that lies at the origin of the gravity-based regression model.

This regression model is analogous to the gravitational attraction force in Newtonian mechanics transformed into a linear form appropriate for a regression through a logarithm. A gravitational force between two bodies *i* and *j* is defined as:

$$F_{\{i,j\}} = \gamma \cdot \frac{M_i \cdot M_j}{d_{ij}^2}$$

Taking the log results in the traditional gravity regression model in the economics trade literature, where the "force" is replaced with a measure of economic (or knowledge) flows between the two entities *i* and *j*:

$$\ln F_{ij} = \gamma + \beta_1 \cdot M_i + \beta_2 \cdot M_j + \beta_3 \cdot d_{ij} + \epsilon_{ij}$$

In the context of subsidiary knowledge inheritance, we posit that an innovation produced by an MNC subsidiary is influenced by two gravitational-like forces: a force generated by the MNC headquarters and acting upon the subsidiary and a competing force generated by the subsidiary's host country and acting upon the subsidiary. Figure 1 shows an intuitive schematic of these competing gravitational forces.

[INSERT FIGURE 1 HERE]

Modeling the Subsidiary-Headquarters Relationship

The (knowledge) trading partners in this case are the subsidiary and the headquarters. Following the gravity model, we define two masses: the mass of the headquarters of the firm (m_{HQ}) and the mass of a subsidiary of the firm (m_S). As explained in detail later in the methods section, the mass is defined as the number of patents granted by the United States Patent Office (USPTO) originating from that entity's location within a year. The measure of (knowledge) trade in our case is the count of patent citations. While traditional gravity models define the distance between the trading partners as geographical distance, such a measure does not necessarily apply to knowledge flows, and thus we introduce a new measurement of knowledge distance based on patent citations, as described in the methods section. In this knowledge flow measure, a flow from an entity to the subsidiary is evidenced through the subsidiary patents' citations of the patents of that entity. For every pair (headquarters, subsidiary) and every year, we thus calculate all citations of the patents of the headquarters in the patents of the subsidiary and define the result as our dependent variable. We hypothesize that the knowledge flow relationship between an MNC headquarters and a foreign-located subsidiary follows a gravitational model. Specifically, we make the following *baseline* hypotheses:

H1a: Knowledge flows from the headquarters to the subsidiary are positively correlated with the stock of knowledge at the headquarters.

H1b: Knowledge flows from the headquarters to the subsidiary are positively correlated with the stock of knowledge at the subsidiary.

H1c: Knowledge flows from the headquarters to the subsidiary are negatively correlated with the knowledge distance between the headquarters and the subsidiary.

Modeling the Subsidiary-Host Country Relationship

The (knowledge) trading partners in this case are the subsidiary and its host country. We define two masses: the mass of the firm's host country (m_C) and the mass of the firm's subsidiary (m_S). The measure of trade is again citation-based. Specifically, for every pair (host country, subsidiary) and every year, we calculate all citations of host country-originated patents in the subsidiary-originated patents and define the result as our dependent variable. The variables related to mass (as explained in detail later) are defined as the number of patents granted by the United States Patent Office (USPTO) originating from that entity's location within a year. We use the same distance measure models as in the previous case. We thus posit that the relationship between a subsidiary and its host country follows a gravitational model, yielding the following *baseline* hypotheses:

H2a: Knowledge flows from the host country to the subsidiary are positively correlated to the host country's relevant stock of knowledge.

H2b: Knowledge flows from the host country to the subsidiary are positively correlated to the subsidiary's stock of knowledge.

H2c: Knowledge flows from the host country to the subsidiary are negatively correlated to the knowledge distance between the host country and the subsidiary.

Comparison of Knowledge Inheritance from Headquarters and Host Country

The central theoretical tension of our paper relates to comparing knowledge inheritance of a multinational subsidiary from the headquarters and the host country and identifying subsidiary level factors that might tilt the balance of knowledge inheritance, away from the headquarters, towards the host country context. We build on the literature of subsidiary absorptive capacity to theorize our core proposition.

There is a rich literature focused on the construct of absorptive capacity in the strategy literature (Cohen and Levinthal, 1990; Zahra and George, 2002). In their seminal article, Cohen and Levinthal (1990) define absorptive capacity of a firm as a function of its prior related knowledge. In other words, the ability of a firm to evaluate and utilize external knowledge is largely a function of the level of prior related knowledge of the firm. As the authors formalize, absorptive capacity refers to the acquisition and assimilation of external knowledge, which is followed by the exploitation of the external knowledge by the firm. Zahra and George (2002) added a third step between assimilating and exploiting knowledge, i.e. transforming the knowledge.

Building on this literature, scholars of multinational firms evolved the concept of subsidiary absorptive capacity. In two related papers, Minbaeva et al. (2003) and Chang, Gong and Peng (2012) showed the subsidiary absorptive capacity is related to facilitating transfer of knowledge from the headquarters and other parts of the MNC network to the subsidiary. Additionally, both of these papers posit that though subsidiary absorptive capacity is an organizational construct, such absorptive capacity resides within the employees of the subsidiary. Minbaeva et al. (2003) articulate that subsidiary absorptive capacity, contained in the prior knowledge stock and intensity of effort of subsidiary employees facilitates the transfer of knowledge from other parts of the MNC to the focal subsidiary. Similarly, Chang, Gong and Peng (2012) posit that subsidiary absorptive capacity is contained in the knowledge bases and motivation of local employees. In fact, knowledge transfers from the headquarters to the subsidiary is facilitated by interactions between expatriates from the headquarters and local employees from the subsidiary and is only successful if local employees have the ability and motivation to utilize such knowledge.

We build on this literature to posit, that over time, as the stock of subsidiary knowledge increases, there is relatively greater subsidiary absorptive capacity to acquire and assimilate knowledge from the local host country knowledge context, rather than from the headquarters. Our theorizing progresses in two distinct steps. First, we build on the insight from prior literature (Chang, Gong and Peng, 2012; Choudhury 2015) that over time, employees transferred from the headquarters to the subsidiary become a relative minority of the employee base in the subsidiary, with the majority comprising local employees. In the second step, we return to the roots of the construct of absorptive capacity. As Cohen and Levinthal (1990) repeatedly stress, this construct is a function of prior related knowledge. As the authors argue, learning at an individual level is cumulative, and learning is easier when the object of learning is related to what the individual already knows. We build on these insights and theorize that local employees, hired from the host country labor market, are more likely to be familiar with knowledge that is embedded in the host country, rather than knowledge that is embedded within the headquarters. As the number of local employees hired by the subsidiary increases over time, we expect to see two effects: the stock of subsidiary knowledge production is likely to increase and local employees are likely to emerge as the dominant group within larger subsidiaries. Additionally, given the greater prior familiarity of local employees with knowledge embedded in the local host country context (rather than knowledge embedded within the headquarters), larger subsidiaries are likely to absorb more knowledge more from the host country, compared to absorbing knowledge from the headquarters. This leads to our final hypothesis:

H3: Larger subsidiaries, characterized by higher knowledge stocks, are likely to absorb relatively more knowledge from the local host country context, compared to absorbing knowledge from the headquarters.

DATA AND METHODS

Sample

Our target data consist of a list of the patents issued to the top 25 US-headquartered patentees as measured by the volume of patents issued at the USPTO. We created a unique dataset of the USPTO patents filed by all of the subsidiaries of these MNCs from 2005 to 2011 (inclusive). The information collected comprised all patent bibliographic information and patent citation data. Because the readily available existing patent datasets did not disambiguate the locations and assignees or the patent textual data for recent years, we created a custom dataset. Our raw dataset comprises all patents issued by the USPTO between the dates of January 1st, 2005 and December 31st, 2011, yielding 1.27 million patents and 69.3 million citations. We extracted this data from Thompson Innovation and enhanced it with string processing techniques to disambiguate company names and inventor locations. The USPTO defines the origin of a patent as the location of residence of the first inventor.[7] Following the same definition, we took additional processing steps to improve the accuracy of the inventor location data.[8]

Variables

For the knowledge inheritance relationship between subsidiary and headquarters, the dependent variable is citations from subsidiary to headquarters within the same company; the independent variables are a measure of the patent output of the subsidiary per year (mass of subsidiary), a measure of the patent output of the headquarters (mass of headquarters), and a measure of the knowledge distance between the innovation outputs of the subsidiary and the headquarters. For the knowledge relationship between subsidiary and host country, the dependent variable is the subsidiary's citations of host country patents, the mass of the headquarters is replaced with the mass of the country, and the knowledge distance is the distance between the innovation outputs of the subsidiary and the outputs of its host country.

Specifications

Subsidiary-headquarters relationship. Our model is a gravity-like model where the dependent variable is the count of a subsidiary's citations of patents filed by the company's headquarters, and our independent variables are, as in a standard gravity model, mass of subsidiary, mass of headquarters, and distance between subsidiary and headquarters. The regression for the case of the subsidiary-headquarters relationship using distances is as follows:

$$BK_{CITES_{ijt}} = \beta_0 + \beta_1 \cdot m_{ijt} + \beta_2 \cdot m_{i_{HQ}t} + \beta_3 \cdot d_{ijt} + \beta_4 \cdot YEAR_t + \beta_5 \cdot FIRM_i \quad (1)$$

where m_{ijt} represents the log of the mass of the subsidiary *j* of company *i* in year *t*, $m_{i_{HQ}t}$ represents the log of the mass of the headquarters of company *i* in year *t*, and d_{ijt} represents the log of the

7U.S. PATENT AND TRADEMARK OFFICE Patent Technology Monitoring Team (PTMT), "Patenting By Geographic Region (State and Country) Breakout By Organization." Accessed 1/4/16. http://www.uspto.gov/web/offices/ac/ido/oeip/taf/stcasg/inx_stcorg.htm

8 Our dataset was stored in a SQL database and further processed with a custom built 2000-line C# program to construct all of the variables in our proposed gravity model.

knowledge distance between subsidiary *j* and the headquarters of company *i* in year *t*, and YEAR and FIRM are fixed effects dummies (all mass variables are stocks of patents). Another equation is used in which similarity measures replace the distance variable above. Due to the limitations of our data, we are unable to distinguish different subsidiaries of the same firm in the same country and therefore consider all R&D activity originating from a firm in a given country to be coming from one subsidiary.

While traditional gravity models define the distance between the trading partners as geographical distance, such a measure does not necessarily apply to knowledge flows. In a typical trade model, geographic distance matters (because of transport costs, border crossing costs, etc.), and in consequence the literature uses the physical distance between the trading partners. Although any citation involves some search cost, the advent of online search engines means that the cost does not depend on geographical distance, discounting the web lag time in accessing the main patent aggregators, such as the European Patent Office or the USPTO, from various international locations. We therefore propose a non-geographic measure of distance that is based on knowledge similarity between the trading parties. Specifically, our gravity model is based on knowledge distance, as in (1), and includes an additional specification based on knowledge similarity measures, shown in (2) below. Both distances and similarity measures are popular in data mining. Model (2), based on a similarity measure σ_{ijt} between subsidiary and headquarters, is:

$$BK_{CITES_{ijt}} = \beta_0 + \beta_1 \cdot m_{ijt} + \beta_2 \cdot m_{i_{HQ}t} + \beta_3 \cdot \sigma_{ijt} + \beta_4 \cdot YEAR_t + \beta_5 \cdot FIRM_i (2)$$

Regarding the dependent variable choice, a traditional gravity model applied to trade would not involve a discrete trade variable. In our case, the back-citation count (our knowledge flow, or trade) is a discrete variable, and therefore a count regression model is more appropriate. We are preserving all other essential aspects of a gravity model (all of the mass and distance variables that remain are logged). In terms of notation, all lower-case variables have already been transformed through a logarithm.

Subsidiary-host country relationship. A gravity model specification for the subsidiary-host country relationship is:

$$BK_{CITES_{ijt}} = \beta_0 + \beta_1 \cdot m_{ijt} + \beta_2 \cdot m_{C_{ijt}} + \beta_3 \cdot d_{ijt} + \beta_4 \cdot YEAR_t + \beta_5 \cdot FIRM_i \ (3),$$

where m_{ijt} represents the log of the mass of the subsidiary *j* of company *i* in year *t*, $m_{C_{ijt}}$ represents the log of the mass of the host country[9] of subsidiary *j* of company *i* in year *t*, d_{ijt} represents the log of the distance between subsidiary *j* of company *i* and its host country in year *t*, and YEAR represents dummies for year fixed effects. The dependent variable is again measured as the subsidiary patents' citations of patents in the host country (not patented by the same firm as the subsidiary).

The model based on a similarity measure[10] σ_{ijt} between a subsidiary and its host country is:

$$BK_{CITES_{ijt}} = \beta_0 + \beta_1 \cdot m_{ijt} + \beta_2 \cdot m_{i_{HQ}t} + \beta_3 \cdot \sigma_{ijt} + \beta_4 \cdot YEAR_t + \beta_5 \cdot FIRM_i \ (4)$$

While we have already defined the masses and the dependent variables for the two relationships of interest (subsidiary-headquarters and subsidiary-host country), we still must define a measure of knowledge distance to properly specify the gravity model.

Measures of Knowledge Distance

Cosine similarity. We implemented several measures of knowledge distance, starting with cosine similarity. The cosine similarity measure is defined as the cosine between two identically-sized vectors. Given vectors $\vec{u}, \vec{v}$, the cosine similarity measure σ is obtained from the dot product of the two vectors:

$$\sigma = \cos(\vec{u}, \vec{v}) = \frac{\sum_{i=1}^{N} u_i \cdot v_i}{\sqrt{\sum_{i=1}^{N} u_i^2} \cdot \sqrt{\sum_{i=1}^{N} v_i^2}}$$

[9] We note that in some cases, the mass of the country as defined in terms of stock of patents is comparable to the mass of the headquarters of some of the firms.

[10] A similarity measure is roughly the equivalent of the inverse of a distance. The mathematical details are explained in the subsection titled "similarity measure to distance".

where $\sigma \in [-1,1]$, $\sigma = -1$ for completely opposite vectors (angle of 180^{O}), $\sigma = 0$ for orthogonal vectors, and $\sigma = 1$ for identical vectors. The cosine similarity measure has been used in the innovation literature (Kay, Newman, Youtie, Porter, & Rafols, 2014) and is widely used in the fields of mathematics and computer science and is one of the most popular similarity measures. Cosine similarity is the measure used in our baseline model.

We use the cosine similarity as follows: for every year, within each firm, we create a vector of patent class counts for every subsidiary and headquarters representing all USPTO patent classes for utility patents (slightly over 400 classes). We weigh the patent class counts based on the total number of patents issued to the subsidiary in that year. (Comparing the raw patent counts per patent class between the subsidiary and the headquarters is not appropriate without taking into account the different total patent outputs of the two entities.) The k^{th} element of the vector of the weighted patent class counts for company *i,* subsidiary *j*, year *t*, and patent class *k* is computed as follows:

$$v_{ijtk} = \frac{Count(issued\ patents\ company\ i\ subsidiary\ j\ year\ t\ in\ class\ k)}{\sum_{k=1}^{M} Count(issued\ patents\ company\ i\ subsidiary\ j\ year\ t\ in\ class\ k)}$$

The patent classes are not sequential. They range from a class of 2 to a class of 987, but with gaps the total number of patent classes is slightly over 400, so *M* in the above is about 400. In the case of our model, patent counts are never negative numbers, so the interval for our similarity measure is $\sigma \in [0,1]$.

This measure is suitable for use along with the previously defined distances because subsidiaries generally follow the research agenda of the headquarters. Specifically, the subdomains of R&D found in the subsidiary's patents are typically a subset of the subdomains of the headquarters, as measured in patents issued per class of the subsidiary and headquarters. We ran regressions (2) and (4) with cosine similarity and reported the results as the baseline model, labeled **model 1.** As this is a similarity measure, the expected coefficient should be positive.

Similarity measure to distance function. A distance function can be intuitively thought of as the inverse of a similarity measure, specifically any transformation *d* satisfying $d(x,x) = 0$, $d(x,y) = d(y,x)$, $d(x,y) = 0 \Leftrightarrow x = y$, and $d(x,z) \leq d(x,y) + d(y,z)$. We used three standard transformations from a similarity measure to a distance function for additional models (labeled as models 2–4 both below and in the results tables):

$$d_1 = -\log(\sigma) \textbf{ (2)}; d_2 = \log(1000 - 1000 \cdot \sigma) \textbf{ (3)}; d_3 = \log\left(1000 - 1000 \cdot \sqrt{\sigma}\right) \textbf{ (4)},$$

where $d_1$is simply a standard irrespective of the minimum value of σ, and d_2 and d_3 represent two additional transformations from a similarity measure to a distance function, taking into account that the lowest non-zero value of similarity in our dataset is of the order of 0.001.[11] These three distances are isomorphic with the cosine similarity in (1), do not change the significance of estimates, and are constructed to place our model into a standard gravity equation, which generally uses a measure of distance between entities.

Bhattacharya coefficient – fidelity similarity. We also used the similarity measure known as the Bhattacharya coefficient, or "fidelity similarity" (Deza & Deza, 2015). For two vectors, this is defined as:

$$\rho(\vec{u}, \vec{v}) = \sum_{i=1}^{N} \sqrt{u_i} \cdot \sqrt{v_i}$$

The fidelity similarity works better than the cosine similarity for vectors with components that are close together; it results in a more compact interval. The fidelity similarity is depicted as model 5 in both the subsidiary to headquarters regression models and the subsidiary to host country regression models.

[11] In our balanced panel dataset covering 25 companies over years 2005-2011, we include all countries where patenting activity occurs. Roughly 50% of the data points contain a zero-patenting subsidiary (thus mass of subsidiary is 0), which implies that the patent counts vector for those subsidiaries is null, resulting in a similarity value of 0.

RESULTS

Summary Statistics

Table 1 represents the summary statistics for the variables in the subsidiary to headquarters regressions (five models, with models 1 and 5 based on similarity measures, and models 2-4 based on distance measures). Table 2 represents the summary statistics for the variables in the subsidiary to host country regressions, following the same five models as in the headquarters to subsidiary case, and sharing one variable with Table 1 (logged mass of subsidiary). Notice that in Tables 1 and 2, the mass measures are plausible (logarithmic scale): the subsidiaries are orders of magnitude smaller than the headquarters, whereas the largest countries are a few orders of magnitude larger than the headquarters (with the country maximum being about 100 times larger than the company maximum). The similarity between the vectors of patenting of headquarters and subsidiaries is typically larger than the similarity between subsidiaries and host countries, which is to be expected.

[INSERT TABLES 1 AND 2 HERE]

The results for the base models are reported in two separate sections, corresponding to the relationship between the headquarters and the subsidiary and the relationship between the host country and the subsidiary. All results include firm and year fixed effects and robust clustered standard errors. The base model results are shown in Table 3 (headquarters-subsidiary) and Table 4 (host country-subsidiary), which report clusters based on country. Table 5 includes robustness checks for the base gravity model for headquarters-subsidiary, while Table 6 includes robustness checks for the base gravity model for host country-subsidiary.

Knowledge Flows between Subsidiary and Headquarters

We find that baseline hypotheses *1a-1c*, which corresponds to a gravitational model for the relationship between an MNC's headquarters and its subsidiary, are validated. All coefficients corresponding to the independent variables are significant and of the expected sign (positive and significant coefficients for masses and negative and significant coefficients for the three distance models, labeled models 2, 3, and 4). Additionally, we introduce a gravity-like model with two similarity measures, cosine similarity (model 1) and fidelity similarity (model 5), which measure affinity in interests between pairs of (headquarters, subsidiary). These similarity measures are also highly significant, and, as expected, have positive coefficients. The results from models 1-5, corresponding to the subsidiary to headquarters relationship, are shown in Table 3. We observed a stronger effect of the mass of the subsidiary (three to five times greater, depending on the model) as compared to the effect of the mass of the headquarters on the measure of knowledge flow from subsidiary to headquarters (the coefficient for the logged headquarters mass ranges from 0.29 to 0.35, whereas the coefficient for the subsidiary mass ranges from 0.67 to 0.81). The closest result to a true gravity relationship is that of the fidelity similarity (column 5), where the coefficient approaches 2. Recall that in classical physics, gravity is modeled as inversely proportional to the square of the distance, a relationship most closely approximated by the fidelity similarity. The cosine similarity is also a good candidate, as the coefficient is between 1 and 2. The similarity measures are highly significant. While cosine similarity is known to the management literature (Younge & Kuhn, 2015), these results show that the fidelity similarity may also be a good candidate for future research on knowledge flows and innovation.

[INSERT TABLE 3 HERE]

Our panel dataset consists mostly of countries that are small in terms of patenting output. Consequently, the panel dataset contains a large number of zeroes for the mass of the subsidiary, which yields zeroes in the citations-dependent variable for the same values of i, j, and t (as a subsidiary with zero patenting activity in a given year does not produce any citations of the headquarters). Considering the literature on specifications for gravity models (Silva & Tenreiro, 2006) and our data, we found a zero-inflated regression model to be most appropriate for this problem. Our dependent variable is a count variable; the most appropriate model is a zero-inflated negative binomial model. We tested binomial versus zero-inflated negative binomial models and found the latter more appropriate.

We excluded three companies with minimal or no international patenting activity from our sample. Specifically, Amazon.com and VERIZON exhibited virtually no foreign subsidiary patents filed with the USPTO during our sample period (2005–2011). The third company removed was AT&T, which has a very complex set of LLCs set up to hold and obfuscate its IP ownership. Because of the low visibility of its IP activities, we were unable to obtain a complete dataset pertaining to AT&T and had to remove it from our regressions.

Knowledge Flows between Subsidiary and Host Country

We find that baseline hypotheses *2a-2c*, corresponding to a gravitational model for the relationship between an MNC subsidiary and its host country, are validated. The results for this section are reported in Table 4. All coefficients corresponding to the independent variables are significant and of the expected sign (positive and significant coefficients for masses and negative and significant coefficients for the three distance models, labeled models 2, 3, and 4). Unlike the case of the subsidiary to headquarters relationship, in the subsidiary-host country relationship, we observe a much stronger effect of the country mass on the knowledge flow as compared to the effect of the subsidiary mass. In the host country-subsidiary relationship, we find the impact of the mass of the host country to be about the same as that of the mass of the subsidiary (the coefficients for both are in the 0.7–0.8 range). This suggests that the relative impact of the host country on the patenting of an R&D subsidiary is stronger than that of the subsidiary's headquarters. We explore this finding further in the marginal effects discussion later in the paper.

[INSERT TABLE 4 HERE]

We also notice that some of the distance/similarity measures perform slightly worse in terms of significance than the same measures in the headquarters-subsidiary case. This is intuitive – the innovation output on average for the countries is much larger than that of the company headquarters, and our knowledge similarity and distance measures are all based on comparing the spectrum of innovation of the two parties (headquarters-subsidiary or subsidiary-host country). If the host country is far larger than the headquarters, it produces diverse innovation and an innovation spectrum that is far noisier than a company subsidiary spectrum; the similarity in such a case is fairly poor. All similarity/distance measures remain significant and of the expected sign. The coefficient on fidelity similarity is closer to 2 than that of the cosine similarity, again suggesting that it is a better fit for a true gravity relationship.

Robustness Checks for Gravity Model

The results are robust to the removal of the large patenting countries, again suggesting that the results are driven by the smaller patenting countries. In addition to the ten models in Tables 2 and 3 corresponding to the five measures of knowledge distance applied to each setting (headquarters-subsidiary and host country-subsidiary), an additional 27 regressions were run to test the robustness of the knowledge flow gravity model. Control variables were obtained from a diverse array of sources, ranging from US State Department visa data to UN immigration data, SCOPUS-based measurements of scientific output, Organisation de Coopération et Devéloppement Economiques (OECD) country-level controls for employment and educational achievement, and the World Bank. The direction and overall magnitude of the main model variables (mass of headquarters, mass of subsidiary, mass of host country, and cosine similarity) were maintained throughout these checks.

The first set of checks, in Table 5, corresponds to four models obtained by adding controls to the base model for the headquarters-subsidiary relationship. Because of the large number of countries (66) in our dataset and the different coverage of those countries in our various data sources, running all of the control variables in one model would result in a vanishingly small subset of our data. To mitigate this problem, we chose to run groups of control variables from the same source within the same model and to split the control variables among four different models. The first column of Table 5 represents the base model for the headquarters-subsidiary relationship (main model independent variables: mass headquarters, mass subsidiary, cosine similarity, with firm and year fixed effects) and is used as a comparison for the next four columns. Model (2) in Table 5 corresponds to the base headquarters-subsidiary model and includes the controls sourced from the OECD.[12] We used two measures of the potential for R&D among a country's population: percent of population attaining tertiary education levels and percent of population attaining a PhD. We chose these measures because highly skilled labor may affect an MNC's decision to locate an R&D subsidiary in a given country and its decision to hire locally, which may in turn affect the knowledge flows to the subsidiary from the headquarters or the host. Yearly data were available for tertiary education; data for PhD graduates, however, were sparser, so we estimated the yearly values using Compound Annual Growth Rate. These two variables were run as part of our "OECD Controls" category and did not affect the results for our main variables; furthermore, PhDs as a percentage of the population were not significant.

SCOPUS is owned by Elsevier and marketed as "the largest database of peer-reviewed literature."[13] It can be used to derive measures of scientific output and quality at the country level. SCImago Journal & Country Rank is a database of country-level measures derived from SCOPUS and made available by SCImago Lab (in partnership with Elsevier).[14] Scientific output (number of articles) and quality (H-index) are measured at the country-year level and are made available by SCImago. We used the

---

[12] OECD Research and Development Indicators; see for example https://data.oecd.org/rd/researchers.htm#indicator-chart, Accessed 1/1/2017.
[13] SCOPUS, https://www.elsevier.com/solutions/scopus. Accessed 1/1/2017.
[14] SCImago Journal & Country Rank, http://www.scimagojr.com/countryrank.php. Accessed 1/1/2017.

country-level variables as controls in Model (3) of Table 5. The overall magnitude of the coefficients of the main variables and sign did not change. Neither control variable was significant.

The US State Department, as a taxpayer-funded agency, makes data of public interest freely available. The Non-Immigrant Visa count per country-year is a relevant source for this study, as certain visa categories are tied to specific types of economic activity. We looked at the following visa categories: B1, one of the most common visa types, which allows for non-immigrant business travel (short stays, such as brief collaborations or conferences); H1B, the most-used skilled worker immigration visa; L1, specifically for intra-company transferees from outside the US to the US and tailored to employees transferring within MNCs; J1, tailored to academic exchanges and used by teachers, scholars, students, and specialists; and O1, extraordinary ability visas reserved for the most desirable specialists and researchers. These types of visas all favor economic exchange between the US and another country and so may be relevant to the headquarters-subsidiary relationship. None of the visa-based variables were significant, and the overall magnitudes and signs of our main coefficients did not change (Model 4, Table 5). Similarly, we used overall immigration counts to the US from other countries as found in UN data[15] (log of number of immigrants per country-year) and found that this variable was not significant. Moreover, the relative magnitudes and directions of the coefficients for the main variables did not change.

We also tested our results using a control for the number of researchers per million inhabitants (sourced from the World Bank) as another measure of a country's potential for R&D. This control variable was not significant and is not reported in Table 5 due to space considerations. Additional robustness checks for the headquarters-subsidiary side include clustering standard errors based on firms (no change, all five models) and, to verify that our effects are not driven by a few countries that

---

[15] UN Population Division compiles migration flows data at the country-year level, http://www.un.org/en/development/desa/population/migration/data/empirical2/migrationflows.shtml. Accessed 1/1/2017.

are very prolific in terms of patenting (such as the UK, India, China, and France), we ran the five models using small patenting countries only. Again, we observed no change in the results. The latter are not reported here because of space limitations.

The checks for the host country-subsidiary relationships are presented in Table 6. As in Table 5, the base model is the first column in Table 6, serving as a reference for the other models. The first set of controls, in column (2) of Table 6, represents the base gravity model of the host country-subsidiary but includes the OECD controls described previously. The third column (Model (3)) of Table 6 includes the SCOPUS-based controls, the fourth column represents results that include the World Bank-sourced number of researchers per million, the fifth column shows results that include the Immigration to US from host country variable (sourced from the UN), and the sixth column of Table 6 includes the State Department Visa issuance counts for categories related to business/worker exchange/skilled labor (B1, J1, L1, H1B, O1).

As in Table 5, all results in Table 6 are based on models that include firm and year fixed effects and clustered robust standard errors (country-level clusters). The magnitudes of the main coefficients and the directions of the effects did not change. In addition to the results reported in Table 6, we ran all five models from Table 4 with firm-level clusters. The results (not reported here due to space limitations) did not change.

[INSERT TABLES 5 AND 6 HERE]

Comparison of Knowledge Inheritance from Headquarters and Host Country Context

H3, a core theoretical proposition of the paper, had stated that larger subsidiaries, characterized by higher knowledge stocks are likely to absorb relatively more knowledge from the local host country context, compared to absorbing knowledge from the headquarters. In other words, we would observe a positive and significant relation between the mass of the subsidiary and the ratio of citations of host country to citations to HQ, as this would indicate a shift in importance of citations to the host country over citations to the firm HQ.[16]

Specifications involving this ratio as the dependent variable are tested in Table 7 (models 1-3) and validate H3. These results indicate a positive and significant relation between subsidiary mass and the ratio of citations to the host country over citations to HQ. To provide further evidence in support of H3, the ratio of the cosine similarity of subsidiary patents to country to the cosine similarity of the subsidiary patents to the HQ is included as a separate DV in models 4-6. We find that the mass of the subsidiary has a highly significant positive relation to the ratio of cosine similarities, which indicates that there is a shift in towards similarity of knowledge of the subsidiary to the host country as opposed to the headquarters for an increase in subsidiary size.

Our results indicate that the host country context is more important for subsidiaries with a growing stock of patented innovations. These results are consistent with the theoretical predictions, as an increase in subsidiary size involves higher absorptive capacity. Large subsidiaries plausibly employ more talent locally, which arguably leads to an increase in the ability to absorb local knowledge, thus yielding a patenting output more similar to the host country. The results in Table 7 are supplemented by a mini-case study on Cisco India, the firm in our sample which converted one of its subsidiaries into a second headquarters.

[INSERT TABLE 7 HERE]

Cisco Second Global Headquarters Case Study. We additionally present a qualitative mini-case study to exemplify our core proposition that larger subsidiaries, characterized by higher knowledge stocks are likely to absorb relatively more knowledge from the local host country context, compared to absorbing knowledge from the headquarters. In the year 2006, the multinational Cisco made an

[16] The dependent variable takes the form of log(1+$\frac{citations\ to\ host\ country}{citations\ to\ HQ}$) to account for the problem of zero citations to host country.

ambitious announcement of setting up a "second global headquarters" in Bangalore, India (Kapur, 2006). At this point, Cisco already had a strong local presence in India and the Cisco Indian workforce comprised more than 1,800 Indian engineers working at the Bangalore R&D facility of Cisco.

The announcement also led to an ambitious innovation mandate awarded to Cisco India, its "largest global development center outside the US" to "develop disruptive business models for Cisco to create new go-to-market channels, markets, processes and technologies for emerging markets."[17] This development awarded the new Global Development Center in Bangalore significant autonomy to develop innovation. At this point Cisco India clearly had a choice of which knowledge pool to tap in order to develop this new innovation.

Based on our theoretical proposition, and especially given the large number of local Indian engineers working at Cisco Bangalore (who arguably had greater familiarity and absorptive capacity to inherit knowledge from the Indian knowledge context compared to inheriting knowledge from the original U.S. based HQ), we expected this event to yield stronger knowledge inflows from the local context innovation to the Cisco India Global Development Center, compared to knowledge inflows from the U.S. Using the framework of the gravity model, we could also expect the knowledge distance between Cisco India and the average Indian patenting entity should shrink post-2006. In other words, an increased knowledge inheritance from India post setting up of the Bangalore global headquarters and the resulting independence from the U.S. HQ should lead to a shrinking knowledge distance between Cisco India and the Indian host country context, post 2006. Table 8 shows the cosine similarity between the patenting in Cisco India and the host country patenting and compares it to the average cosine similarity of all firms in our sample to patenting in India; there is a jump in similarity between Cisco India and its host country post-2006 as compared to the average Indian subsidiary in our sample. Further, we see the similarity of the Cisco Indian patenting to its host country increases

[17] Cisco Company Overview, https://www.cisco.com/c/en_in/about/company-overview.html. Accessed 3/26/2018.

year over year, as theory predicts for a subsidiary growing annually in local headcount. Cisco India additionally increased the hiring of local talent dramatically beginning with 1,800 employees in 2006 (Kapur, 2006), growing rapidly to 2,800 in 2007[18] , and reaching 11,800 in FY'13 (source: Cisco report[19]). Arguably, this illustrates our arguments about how an increase in hiring local talent could result in an increase in the ability of the firm to absorb local knowledge from the host country.

[INSERT TABLE 8 HERE]

A second example, providing similar evidence, relates to IBM India. In 2007, IBM initiated a strategy of disproportionately hiring its R&D workforce in India (Goel, 2017), and, while not designating the subsidiary a second HQ formally, currently employs more talent in India compared to the U.S. (Goel 2017). The advantages of a highly qualified labor pool at a fraction of the cost at the headquarters are part of why top US MNCs locate their R&D in India. While less of a formal arrangement than in the case of CISCO, we still see the same trends post-expansion in India, specifically a jump in cosine similarity between IBM India and its host country after 2007, that jump increasing year-to-year faster than the the average Indian subsidiary, as shown in Table 9.

[INSERT TABLE 9 HERE]

These qualitative examples, while limited in scope to the two individual firms, help interpret our findings, and serve as an example of how hiring local talent might shift the balance of subsidiary knowledge inheritance from the to the host country context.

CONCLUSIONS AND DISCUSSION

In this paper, we use the gravity equation in economics to estimate two comparable specifications – one measuring knowledge flows from the headquarters to a multinational subsidiary and the second

---

18 Cisco to Double Headcount in India, FirstPost, https://www.firstpost.com/business/biztech/cisco-to-double-headcount-in-india-1860809.html, Accessed 7/8/2018.

19 Cisco India Overview, http://www.audentia-gestion.fr/cisco/pdf/Cisco_India_Overview_New.pdf, Accessed 7/1/2018. Annual reports for Cisco usually contain only overall Asia-Pacific employee headcounts; India specific headcounts are only available for a few years, as in this report. Secondary sources such as news articles have been used to obtain headcount data for some years.

measuring knowledge flows from the host country context to the subsidiary. Our empirical apparatus allows us to consider a relatively robust "apples to apples" comparison of these two knowledge flows. Using unique data on patent citations and the "masses" of patents filed with the USPTO by the headquarters, subsidiaries and host countries of the top 25 US-headquartered MNCs from 2005 to 2011, we validate the gravity specification for both headquarters to subsidiary knowledge flows and host country to subsidiary knowledge flows. Our results indicate that the role of the subsidiary mass is several times more important than the role of the headquarters mass. Moreover, the influence of the subsidiary mass on the knowledge flow is increasing faster than proportionally. This is a departure from the standard gravitational model. We also compare the *relative influence* of the headquarters and the host country context on knowledge flows into the MNC subsidiary. Our results indicate that as the size of the subsidiary increases, the host country's influence on knowledge inflows into the subsidiary grows faster than the influence of the headquarters. In other words, MNC subsidiaries may differ in the extent to which they are influenced by the host country as compared to the headquarters. Specifically, subsidiaries that have a greater stock of patented innovations may be more susceptible to the host country's influence with reference to knowledge inheritance.

Contributions

Our study contributes to several streams of the strategy and IB literature focused on multinational firms. Our results contribute to the literature on how multinational subsidiaries evolve their capabilities and charter over time. Birkinshaw and Hood (1998) define subsidiary evolution as the enhancement or depletion of the capabilities of the subsidiary, a process that is affected by host country factors such as relative cost of factor inputs, the "strategic importance" of the host country, host country government support and "dynamism" of the local business environment. Given this framing, the authors outline several pathways to how multinational subsidiaries might evolve their charter over time. One of these paths relate to "subsidiary-driven charter extension" (SDE), which involves a "long and often slow process" where subsidiary managers search for new market opportunities and develop capabilities to fulfill this opportunity. However, in the SDE evolution process, the final step involves subsidiary managers pitching their proposal for pursuing this opportunity to the headquarters. As the authors state, MNC subsidiaries "seek out and develop new business opportunities and then put them forward to parent company managers" (Birkinshaw and Hood, 1998; page 1998). Our findings suggest that the literature on multinationals needs to view the subsidiary-host-country-headquarters triadic relationship in new light. It is possible that MNC subsidiaries might endogenously shift their center of gravity of knowledge inheritance from the headquarters to the host country. Though our study is restricted to analyzing conditions under which the host country (and not the HQ) becomes the center of gravity of knowledge inheritance for the MNC subsidiary, it is plausible that this shifting center of gravity might affect governance and decision making incentives for subsidiary managers. Under these conditions, future researchers should study whether the "subsidiary-driven charter extension" might involve bypassing approval seeking from the headquarters. Future research in this area should study conditions under which subsidiaries might independently (i.e. without approval of the headquarters) seek to change their charter and capabilities.

Our results respond to the recent call in the strategy and international business literature for firms to develop contextual intelligence in host countries (Dhanaraj & Khanna, 2011; Khanna, 2015). Meyer, Mudambi, and Narula (2011) predict that host countries will play an increasingly important role in shaping MNC subsidiaries and thus MNCs overall. Santos and Williamson (2015) advise MNCs to cultivate a local presence that is not merely "adaptive" but fully intertwined with or even "made" in the local context. One way to establish a local presence is by learning from the host country context. In summary, our results represent a step forward in empirically measuring reverse innovation (Govindrajan & Ramamurti, 2011) and comparing knowledge flows from the host country to the subsidiary. Our study also contributes to the recent call in the IB literature to view the host country context in a more "positivist" light, rather than using a lens of cultural distance and the liability of foreignness. In fact, the recent MNC literature has urged scholars to think beyond the "liability of foreignness" lens, and view the host country in more positive light given that MNCs can exploit the host country knowledge pool to "develop unique and potentially valuable capabilities, and foster learning and innovation" (Stahl, Tung, Kostova, & Zellmer-Bruhn, 2016, p. 623). Our results also contribute to the literature on knowledge flows within MNCs by empirically measuring knowledge inflows into the subsidiary from the host country context.[20]

In addition to contributing to the MNC literature, we contribute to the knowledge flow and gravity literatures by introducing unique measures of knowledge distance. While we introduce the gravity model to the strategy and innovation literature, we depart from the traditional gravity model that uses a physical distance measure and instead propose a measurement of knowledge distance based on cosine similarity and several transformations of the cosine similarity, and we introduce fidelity similarity (the Bhattacharya coefficient) to the management literature. We find the fidelity similarity to be the truest to the classical physics gravity model and propose it for future use in our literature. The similarity approach presented here is useful for comparing the patenting outputs of entities in general and is not limited to MNCs; in other words, we expect this part of our study to be generalizable.

Limitations and Future Research Directions

We are aware that our study is limited by our use of only US-headquartered MNCs and USPTO data. Even with this limitation, many errors in firm names in the raw patent data had to be manually corrected. Future studies should examine the influences of the communication and control forms

[20] The literature on knowledge flows within MNCs has used several perspectives, including using knowledge communication and transmission theory (Gupta & Govindrajan, 2000); network theory (Ghoshal & Bartlett, 1990; Hansen, 2002); cluster innovation (Alcacer & Zhao, 2012); institutional theory (Kostova & Roth, 2002); modularity (Zhao, 2006) and theories of human capital mobility (Almeida & Kogut, 1999; Choudhury, 2016; Oettl & Agrawal, 2008; Song, Almeida, & Wu, 2003, etc.). However, as stated earlier, this literature has mostly ignored knowledge inflows from the host country.

practiced by MNC headquarters, as identified by Nobel and Birkinshaw (1998), on the knowledge flow among subsidiaries, the headquarters, and the host countries. Finally, a detailed mathematical analysis of the distances used in this study might explain why, while some of them are isomorphic (d_1, d_2, d_3), they are considerably different in the extent to which they reveal differences in innovative processes between entities. We suggest that the methodology of evaluating distances by the spectrum of the interest in the innovative domains, as proposed in this study and found outstandingly effective, could be applicable to a broader class of economic statistical models. In future work, this methodology deserves to be tested in various other problems of knowledge transfer.

Conclusion

We began this study with a question of whether there could be a shifting center of gravity of subsidiary knowledge inheritance from the headquarters to the host country context. Our study sheds light on this question and has implications for the literature of multinationals in light of the recent protectionist trends across developed and emerging markets of the world. If the phenomenon of skepticism towards the role of headquarters, de-globalization or reversal of globalization persists, it would accentuate the importance of the local contexts to multinational subsidiaries.[21] Our aim is *not* to undermine the importance of the headquarters to the MNC subsidiary on the question of knowledge inheritance. Our attempt is merely to shift the needle in scholarship and achieve a more balanced perspective on how MNC subsidiaries can inherit knowledge from both its parents (i.e. the HQ and the host), rather than only from the HQ.

[21] The phenomenon of de-globalization has a deep historical precedence and extant research has established that MNC subsidiaries can flourish in these historical times by being more firmly tethered to the host country context (Choudhury and Khanna, 2014).

REFERENCES

Agrawal, A., Cockburn, I., & McHale, J. 2006. Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. ***Journal of Economic Geography***, 6(5): 571-591.

Agrawal, A., Kapur, D., & McHale, J. 2008. How do spatial and social proximity influence knowledge flows? Evidence from patent data. ***Journal of Urban Economics***, 64(2): 258-269.

Alcacer, J., & Zhao, M. 2012. Local R&D strategies and multilocation firms: The role of internal linkages. ***Management Science,*** 58(4): 734–753.

Almeida, P., & Kogut, B. 1999. Localization of knowledge and the mobility of engineers in regional networks. ***Management Science,*** 45(7): 905–917.

Almeida, P., & Phene, A. 2004. Subsidiaries and knowledge creation: The influence of the MNC and host country on innovation. ***Strategic Management Journal,*** 25(8-9): 847–864.

Anand, J., & Delios, A. 2002. Absolute and relative resources as determinants of international acquisitions. ***Strategic Management Journal***, *23*(2): 119-134.

Anderson, J.E. 1979. A theoretical foundation for the gravity equation. ***The American Economic Review***, 69(1): 106-116.

Andersson, U., Forsgren, M., & Holm, U. 2002. The strategic impact of external networks: Subsidiary performance and competence development in the multinational corporation. ***Strategic Management Journal,*** 23(11): 979–996.

Anderson, J.E. 1979. A theoretical foundation for the gravity equation. ***The American Economic Review***, 69(1): 106-116.

Anderson, J. E., & van Wincoop, E. 2003. Gravity with gravitas: A solution to the border puzzle. ***American Economic Review,*** 93(1): 170–192.

Baier, S.L. and Bergstrand, J.H. 2009. Estimating the effects of free trade agreements on international trade flows using matching econometrics. ***Journal of International Economics***, 77(1): 63-76.

Bartlett, C. A. 1986. Building and managing the transnational: The new organizational challenge**.** In M. Porter (Ed.), ***Competition in global industries***, pp. 367-401. Boston, MA: Harvard Business School Press.

Bartlett, C. A., Doz, Y., & Hedlund, G. 2012. ***Managing the global firm*** (RLE International Business). Vol. 3. London: Routledge.

Bergstrand, J.H. 1985. The gravity equation in international trade: Some microeconomic foundations and empirical evidence. ***The Review of Economics and Statistics***, 67(3): 474-481.

Bergstrand, J. H. 1989. The generalized gravity equation, monopolistic competition, and the factor-proportions theory in international trade. ***The Review of Economics and Statistics,*** 71(1): 143–153.

Berry, H. 2015. Knowledge inheritance in global industries: The impact of parent firm knowledge on the performance of foreign subsidiaries. ***Academy of Management Journal***, *58*(5): 1438-1458.

Birkinshaw, J., & Hood, N. 1998. Multinational subsidiary evolution: Capability and charter change in foreign-owned subsidiary companies. ***Academy of Management Review***, 23(4): 773-795.

Birkinshaw, J. M., & Morrison, A. J. 1995. Configurations of strategy and structure in subsidiaries of multinational corporations. ***Journal of International Business Studies,*** 26(4): 729–753.

Björkman, I., Barner-Rasmussen, W., & Li, L. 2004. Managing knowledge transfer in MNCs: The impact of headquarters control mechanisms. ***Journal of International Business Studies,*** 35(5): 443–455.

Bouquet, C. and Birkinshaw, J., 2008. Weight versus voice: How foreign subsidiaries gain attention from corporate headquarters. **Academy of Management Journal**, 51(3), pp.577-601.

Buckley, P. J., & Casson, M. 1976. ***The future of the multinational enterprise***. Vol. 1. London: MacMillan.

Cantwell, J., & Mudambi, R. 2005. MNE competence-creating subsidiary mandates. ***Strategic Management Journal***, *26*(12): 1109-1128.

Caves, R. E. 1971. International corporations: The industrial economics of foreign investment. ***Economica,*** 38(149): 1–27.

Caves, R. E. 1974. Multinational firms, competition, and productivity in host-country markets. ***Economica,*** 41(162): 176–193.
Caves, D. W., Christensen, L.R., & Diewert, W. E. 1982. The economic theory of index numbers and the measurement of input, output, and productivity. ***Econometrica: Journal of the Econometric Society,*** 50(6): 1393–1414.
Chacar, A.S., Newburry, W. and Vissa, B., 2010. Bringing institutions into performance persistence research: Exploring the impact of product, financial, and labor market institutions. ***Journal of International Business Studies***, 41(7), pp.1119-1140.
Chang, Y. Y., Gong, Y., & Peng, M. W. 2012. Expatriate knowledge transfer, subsidiary absorptive capacity, and subsidiary performance**. *Academy of Management Journal***, *55*(4): 927-948.
Choudhury, P. 2015. Return migration and geography of innovation in MNEs: a natural experiment of knowledge production by local workers reporting to return migrants**. *Journal of Economic Geography***, *16*(3): 585-610.
Choudhury, P., & Khanna, T. 2014. Charting Dynamic Trajectories: Multinational Enterprises in India. ***Business History Review***, *88*(1): 133-169.
Cohen, W. M., & Levinthal, D. A. 1990. Absorptive capacity: A new perspective on learning and innovation. ***Administrative Science Quarterly,*** 35(1): 128-152.
Dacin, M. T., Beal, B. D., & Ventresca, M. J. 1999. The embeddedness of organizations: Dialogue & directions. **Journal of Management**, 25(3): 317–356.
Deza, M. M., & Deza, E. 2015. ***Encyclopedia of distances***. Berlin: Springer.
Doz, Y. L., Bartlett, C. A., & Prahalad, C. K. 1981. Global competitive pressures and host country demands managing tensions in MNCs. ***California Management Review***, 23(3): 63–74.
Dhanaraj, C., and Khanna, T. 2011. Transforming mental models on emerging markets. ***Academy of Management Learning & Education,*** 10(4): 684–701.
Dunning, J. H. 1998. Location and the multinational enterprise: a neglected factor? ***Journal of International Business Studies***, *29*(1): 45-66.
Feinberg, S. E., & Gupta, A. K. 2004. Knowledge spillovers and the assignment of R&D responsibilities to foreign subsidiaries. ***Strategic Management Journal***, 25: 823–845.
Ghoshal, S. 1986. ***The innovative multinational: A differentiated network of organizational roles and management processes***. Unpublished thesis, Harvard University, Graduate School of Business Administration.
Ghoshal, S. 1987. Global strategy: An organizing framework. ***Strategic Management Journal***, 8(5): 425–440.
Ghoshal, S., & Bartlett, C. A. 1990. The multinational corporation as an interorganizational network. ***Academy of Management Review***, 15(4): 603–626.
Ghoshal, S., & Nohria, N. 1989. Internal differentiation within multinational corporations. ***Strategic Management Journal,*** 10(4): 323–337.
Goel, V. 2017. IBM Now Has More Employees in India Than in the U.S. *The New York Times* (Sept 28), https://www.nytimes.com/2017/09/28/technology/ibm-india.html. Accessed June 14th, 2018.
Govindarajan, V. & Ramamurti, R. 2011. Reverse innovation, emerging markets, and global strategy. ***Global Strategy Journal***, 1(3-4): 191-205.
Gupta, A. K., & Govindarajan, V. 1991. Knowledge flows and the structure of control within multinational corporations. ***Academy of Management Review***, 16(4): 768–792.
Gupta, A. K., & Govindarajan, V. 2000. Knowledge flows within multinational corporations. ***Strategic Management Journal,*** 21(4): 473–496.
Hansen, M. T. 2002. Knowledge networks: Explaining effective knowledge sharing in multiunit companies. ***Organization Science***, 13(3): 232–248.
Hymer, S.H. 1960. ***The International Operations of National Firms: A Study of Direct Investment.*** Unpublished Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge.
Hymer, S. H. 1976. ***The international operations of national firms: A study of direct foreign investment.*** Vol. 14. Cambridge, MA: MIT Press.
Jaffe, A.B., Trajtenberg, M. and Henderson, R. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. ***Quarterly Journal of Economics***, 108(3): 577-598.

Kapur, M. 2006. Cisco to set up second head office in Bangalore. *The Times of India* (Nov 2), https://timesofindia.indiatimes.com/business/india-business/Cisco-to-set-up-second-head-office-in-Blore/articleshow/287171.cms. Accessed April 16th, 2018.
Kay, L., Newman, N., Youtie, J., Porter, A.L., & Rafols, I. 2014. Patent overlay mapping: Visualizing technological distance. ***Journal of the Association for Information Science and Technology,*** 65(12): 2432–2443.
Khanna, T., Palepu, K.G. and Sinha, J., 2005. Strategies that fit emerging markets. ***Harvard Business Review***, *83*(6), pp.4-19.
Khanna, T. & Palepu, K.G., 2010. Winning in emerging markets: A road map for strategy and execution. Harvard Business Press.
Khanna, T. 2015. A case for contextual intelligence. ***Management International Review,*** 55(2): 181–190.
Kindleberger, C. P. 1969. American business abroad. ***The International Executive,*** 11(2): 11–12.
Kostova, T., & Roth, K. 2002. Adoption of an organizational practice by subsidiaries of multinational corporations: Institutional and relational effects. ***Academy of Management Journal,*** 45(1): 215–233.
Kuemmerle, W., 1999. The drivers of foreign direct investment into research and development: an empirical investigation. ***Journal of International Business Studies***, 30(1): 1-24.
Lewer, J. J., & Van den Berg, H. 2008. A gravity model of immigration. ***Economics Letters,*** 99(1): 164–167.
Mátyás, L. 1997. Proper econometric specification of the gravity model. ***The World Economy***, 20(3): 363–368.
Meyer, K. E., & Estrin, S. 2014. Local context and global strategy: extending the integration responsiveness framework to subsidiary strategy. ***Global Strategy Journal***, *4*(1): 1-19.
Meyer, K. E., Mudambi, R., & Narula, R. 2011. Multinational enterprises and local contexts: The opportunities and challenges of multiple embeddedness. ***Journal of Management Studies,*** 48(2): 235–252.
Minbaeva, D., Pedersen, T., Björkman, I., Fey, C. F., & Park, H. J. 2003. MNC knowledge transfer, subsidiary absorptive capacity, and HRM. ***Journal of International Business Studies***, *34*(6): 586-599.
Monteiro, L. F., Arvidsson, N., & Birkinshaw, J. 2008. Knowledge flows within multinational corporations: Explaining subsidiary isolation and its performance implications. ***Organization Science***, 19(1): 90–107.
Mudambi, R., Pedersen, T., & Andersson, U. 2014. How subsidiaries gain power in multinational corporations. ***Journal of World Business***, 49(1): 101–113.
Narula, R. 2014. Exploring the paradox of competence-creating subsidiaries: Balancing bandwidth and dispersion in MNEs. ***Long Range Planning***, 47(1): 4–15.
Nell, P. C., & Ambos, B. 2013. Parenting advantage in the MNC: An embeddedness perspective on the value added by headquarters. ***Strategic Management Journal***, 34(9): 1086–1103.
Nelson, R., & Winter, S. 1982. ***An evolutionary theory of economic change***. Cambridge, MA: Belknap Press of Harvard University Press.
Nobel, R., & Birkinshaw, J. 1998. Innovation in multinational corporations: Control and communication patterns in international R&D firms. ***Strategic Management Journal***, 19: 479–496.
Nohria, N., & Ghoshal, S. 1994. Differentiated fit and shared values: Alternatives for managing headquarters-subsidiary relations. ***Strategic Management Journal***, 15(6): 491–502.
Oettl, A., & Agrawal, A. 2008. International labor mobility and knowledge flow externalities. ***Journal of International Business Studies***, 39(8): 1242-1260.
O'Donnell, S. W. 2000. Managing foreign subsidiaries: Agents of headquarters, or an interdependent network? ***Strategic Management Journal,*** 21(21): 525–548.
Porter, M. E. 1986. ***Competition in global industries***. Boston, MA: Harvard Business Press.
Prahalad, C. K., & Doz, Y. L. 1987. The multinational mission: Balancing global integration with local responsiveness. *New York, NY*.
Rodrik, D., Subramanian, A., & Trebbi, F. 2004. Institutions rule: the primacy of institutions over geography and integration in economic development. ***Journal of Economic Growth***, *9*(2): 131-165.
Rugman, A. M., & Verbeke, A. 1992. A note on the transnational solution and the transaction cost theory of multinational strategic management. ***Journal of International Business Studies,*** 23(4): 761–771.
Rugman, A., & Verbeke, A. 2001. Subsidiary-specific advantages in multinational enterprises. ***Strategic Management Journal***, 22(3): 237–250.

Santos, J.F., & Williamson, P.J. 2015. The new mission for multinationals. ***MIT Sloan Management Review***, 56(4): 45.
Silva, J. M. C. S., & Tenreyro, S. 2006. The log of gravity. ***The Review of Economics and Statistics,*** 88(4): 641–658.
Singh, J. 2007. Asymmetry of knowledge spillovers between MNCs and host country firms. ***Journal of International Business Studies***, 38(5): 764–786.
Singh, J. 2008. Distributed R&D, cross-regional knowledge integration and quality of innovative output. ***Research Policy***, 37(1): 77–96.
Song, J., Almeida, P., & Wu, G. 2003. Learning-by-hiring: When is mobility more likely to facilitate interfirm knowledge transfer? ***Management Science***, 49(4): 351–365.
Stahl, G. K., Tung, R. L., Kostova, T., & Zellmer-Bruhn, M. 2016. Widening the lens: Rethinking distance, diversity, and foreignness in international business research through positive organizational scholarship. ***Journal of International Business Studies,*** 47(6): 621–630.
Summary, R.M. 1989. A political-economic model of US bilateral trade. ***Review of Economics and Statistics***, 71(1):179-182.
Teece, D. J. 1981. The market for know-how and the efficient international transfer of technology. ***The Annals of the American Academy of Political and Social Science***, 458(1): 81–96.
Tinbergen, J. 1962. The world economy. *Suggestions for an International Economic Policy. New York, NY: Twentieth Century Fund.*
Vernon, R. 1966. International investment and international trade in the product cycle. ***Quarterly Journal of Economics,*** 80(2): 190–207.
Waugh, M.E. 2010. International trade and income differences. ***American Economic Review***, 100(5): 2093-2124.
Younge, K. A., & Kuhn, J. M. 2015. Patent-to-patent similarity: A vector space model. Available at SSRN.
Zahra, S. A., & George, G. 2002. Absorptive capacity: A review, reconceptualization, and extension. ***Academy of Management Review,*** 27(2): 185–203.
Zhao, M. 2006. Conducting R&D in countries with weak intellectual property rights protection. ***Management Science***, 52(8): 1185–1199.

Table 1: Summary Statistics for Variables in Subsidiary to HQ Regressions

	(mean)	(sd)	(min)	(max)
Cites Subsidiary to HQ	.9570815	4.547072	0	82
Log(Subsidiary Mass)	.8209908	1.10791	0	5.204007
Log(HQ Mass)	6.414829	1.033538	1.791759	8.483843
Cosine Similarity	.1780619	.2593527	0	1
Distance d_1	4.275926	8.166304	.0145975	6.907755
Distance d_2	6.61518	.5400273	2.673558	6.906755
Distance d_3	6.391382	.7381062	1.984054	6.875622
Fidelity Similarity	.1726259	.2269262	.001	.8693181

Table 2: Summary Statistics for Variables in Subsidiary to Country Data

	(mean)	(sd)	(min)	(max)
Cites Subsidiary to Country	.4325567	2.748737	0	59
Log(Subsidiary Mass)	.8209908	1.10791	0	5.204007
Log(Country Mass)	6.404508	2.364738	0	10.79853
Cosine Similarity	.099099	.1573169	.001	1
Distance d_1	4.509807	2.544627	0	6.907755
Distance d_2	6.782846	.2363022	3.921989	6.906755
Distance d_3	6.603674	.3943458	3.241714	6.875622
Log(Fidelity Similarity)	-4.392618	2.606018	-6.907755	0

Table 3: Subsidiary to Headquarters Regressions - Base Model and Additional Specifications Based on Knowledge Distance

	(1)	(2)	(3)	(4)	(5)
Log(Mass HQ)	0.294^{+}	0.317 $^{+}$	0.294^{+}	0.317^{+}	0.325^{+}
	(0.178)	(0.183)	(0.178)	(0.183)	(.178)
Log(Mass Subsidiary)	0.801***	0.811 ***	0.801***	0.811***	0.679***
	(0.064)	(0.131)	(0.064)	(0.131)	(0.068)
Cosine Similarity	1.301***				
	(0.247)				
Distance d_1		-0.245			
		(0.184)			
Distance d_2			-0.0013***		
			(0.0002)		
Distance d_3				-0.245	
				(0.184)	
Fidelity Similarity					2.079**
					(.392)
_cons	-7.450***	-6.601***	-6.149***	-6.601***	-7.578***
	(1.188)	(1.788)	(1.245)	(1.788)	(1.154)
Year FE	yes	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes	yes
inflate					
log_mass_subsidiary	-4.316**	-2.373	-4.316**	-2.374	-4.328**
	(1.643)	(3.000)	(1.643)	(3.000)	(-3.01)
_cons	3.655***	2.377	3.656***	2.377	3.660***
	(0.681)	(1.865)	(0.681)	(1.864)	(0.622)
lnalpha	-0.354	-0.459	-0.355	-0.321	-0.321
N	3157	3157	3157	3157	3157
Clusters	69	69	69	69	69

Standard Errors in parentheses (country level clustering)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The first model (1) represents the base model for the headquarters to subsidiary relationship. The dependent variable is the count of citations from the subsidiary patents to the headquarter patents. Main model variables represent stock of knowledge at headquarters as represented by patenting output (Mass HQ), stock of knowledge at subsidiary (Mass Subsidiary), and the cosine similarity between knowledge production at headquarters and subsidiary. The following three models, (2) through (4) use the same base specification, except the cosine similarity is replaced by different measures of knowledge distance between subsidiary and headquarters (as described in the paper). The fifth model (column (5)) utilizes the same base specification, except that the cosine similarity is replaced by the fidelity similarity. All five models include firm and year fixed fixed effects as well as robust clustered standard errors (clustered based on country, 69 clusters). Specifications including controls for the Headquarters to Subsidiary base model are in Table 5.

Table 4: Subsidiary to Host Country Regressions Base Model and Additional Specifications Based on Knowledge Distance

	(1)	(2)	(3)	(4)	(5)
Log(Country Mass)	0.757***	0.736***	0.762***	0.759***	0.767***
	(0.450)	(0.039)	(0.047)	(0.046)	(0.048)
Log(Subsidiary Mass)	0.782***	0.816***	0.805***	0.796***	0.675***
	(0.091)	(0.096)	(0.089)	(0.090)	(0.118)
Cosine Similarity	1.306**				
	(0.548)				
Distance d_1		-0.125*			
		(0.109)			
Distance d_2			-0.555$^+$		
			(0.030)		
Distance d_3				-0.441**	
				(0.226)	
Fidelity Similarity					2.357*
					(1.04)
_cons	-13.043***	-12.403***	-9.188***	-10.042***	-13.130***
	(0.681)	(0.685)	(2.116)	(1.610)	(0.733)
Year FE	yes	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes	yes
inflate					
Log(Mass Subsidiary)	-2.080***	-1.907***	-2.120***	-2.081***	-2.100***
	(0.223)	(0.283)	(0.209)	(0.215)	(0.231)
_cons	2.479***	2.301***	2.549***	2.492***	2.500***
	(0.345)	(0.443)	(0.332)	(0.348)	(0.347)
lnalpha	-0.294	-0.288	-0.283	-0.285	-0.284
N	3157	3157	3153	3153	3157
Clusters	69	69	69	69	69

Standard Errors in parentheses (country clusters)
$^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The first model (1) represents the base model for the host country to subsidiary relationship. The dependent variable is the count of citations from the subsidiary patents to the host country patents. Main model variables represent stock of knowledge of the host country as represented by patenting output (Country Mass), stock of knowledge at subsidiary (Mass Subsidiary), and the cosine similarity between knowledge production in the host country and that of the subsidiary. Models (2) through (4) use the same base specification, except the cosine similarity is replaced by different measures of knowledge distance between subsidiary and country (as described in the paper). The fifth model (column (5)) utilizes the same base specification, except that the cosine similarity is replaced by the fidelity similarity. All five models include firm and year fixed fixed effects as well as robust clustered standard errors (clustered based on country, 69 clusters). Specifications including controls for the Country to Subsidiary base model are in Table 6.

Table 5: Robustness Checks - Subsidiary to Headquarters Regressions - Base Model in Comparison with Models with Various Controls

	(1)	(2)	(3)	(4)	(5)
Log(Mass HQ)	0.294^{+}	0.455^{+}	0.223	0.293	0.275
	(0.178)	(0.274)	(0.170)	(0.193)	(0.181)
Log(Mass Subsidiary)	0.801^{***}	0.820^{***}	0.880^{***}	0.821^{***}	0.818^{***}
	(0.064)	(0.084)	(0.079)	(0.075)	(0.067)
Cosine Similarity	1.301^{***}	1.259^{***}	1.206^{***}	1.137^{***}	1.294^{***}
	(0.247)	(0.333)	(0.273)	(0.286)	(0.257)
_cons	-7.450^{***}	-9.540^{***}	-5.326^{***}	-6.751^{***}	-7.284^{***}
	(1.188)	(1.949)	(1.334)	(1.310)	(1.375)
Year FE	yes	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes	yes
OECD Controls		yes			
SCOPUS Controls			yes		
State Dept Controls				yes	
Imm to US Control (UN)					yes
inflate					
log_mass_subsidiary	-4.316^{**}	-7.501^{*}	-4.166^{+}	-4.269^{*}	-4.030^{*}
	(1.643)	(3.600)	(2.299)	(1.674)	(1.954)
_cons	3.655^{***}	5.282^{**}	3.578^{***}	3.798^{***}	3.484^{***}
	(0.681)	(1.928)	(0.919)	(0.701)	(0.777)
lnalpha	-0.354	-0.206	-0.406	-0.462	-0.377
N	3157	1998	3074	2750	3074
Clusters	69	26	68	56	68

Standard Errors in parentheses (country level clustering)
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The first model (1) represents the base model for the headquarters to subsidiary relationship.
The following four models include controls added to the base specification model. The control variables were obtained from different sources with varying coverage on countries, thus groups of control variables from the same source are displayed in different columns as part of separate regressions. Model (2) is the base specification plus OECD controls measuring research characteristics at the country level. Model (3) is the base specification plus controls based on SCOPUS (source: SCImago country data) which measure country level scientific throughput. Model (4) is the base specification plus controls based on State Department Visa issuance by country. Model (5) represents the base specification plus a control c variable for overall immigration to the US based on country of origin (data source: UN). Additional robustness checks were run and are discussed in the paper. All results in this table include firm and year fixed effects as well as robust clustered standard errors (clustered based on country).

Table 6: Robustness Checks - Subsidiary to Host Country Regressions Subsidiary to Headquarters Regressions - Base Model in Comparison with Models with Various Controls

	(1)	(2)	(3)	(4)	(5)	(6)
Log(Cntry Mass)	0.757***	0.772***	0.844***	0.757***	0.710***	0.831***
	(0.450)	(0.045)	(0.053)	(0.043)	(0.041)	(0.214)
Log(Sub. Mass)	0.782***	0.905***	0.963***	0.977***	0.901***	0.816***
	(0.091)	(0.058)	(0.064)	(0.087)	(0.058)	(0.816)
Cosine Sim.	1.306**	1.076+	0.499	1.519*	0.798	1.332*
	(0.548)	(0.643)	(0.572)	(0.446)	(0.588)	(0.582)
_cons	-13.043***	-17.668***	-11.042***	-14.601***	-13.108***	-13.435***
	(0.681)	(1.076)	(1.507)	(1.114)	(0.599)	(0.903)
Year FE	yes	yes	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes	yes	yes
OECD Ctrls		yes				
SCOPUS Ctrls			yes			
Res per Mln				yes		
Imm to US (UN)					yes	
Visas (State)						yes
inflate						
Log(Subs Mass)	-2.080***	-2.106***	-2.096***	-1.979***	-2.149***	-2.096***
	(0.223)	(0.302)	(0.270)	(0.251)	(0.269)	(0.265)
_cons	2.479***	2.142***	2.367***	2.247***	2.419***	2.410***
	(0.345)	(0.357)	(0.401)	(0.314)	(0.417)	(0.341)
lnalpha	-0.294	-0.571	-0.482	-1.116	-0.409	-0.337
N	3157	1998	3074	2397	3080	2750
Clusters	69	26	68	55	68	56

Standard Errors in parentheses (country clusters)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The first model (1) represents the base model for the subsidiary to host country relationship. The following four models include controls added to the base specification model. The control variables were obtained from different sources with varying coverage on countries, thus groups of control variables from the same source are displayed in different columns as part of separate regressions. Model (2) represents the base specification plus OECD controls measuring research characteristics at the country level. Model (3) represents the base specification plus controls based on SCOPUS data (source: SCImago Country level data) which measure country level scientific throughput. Model (4) is the base specification plus a control based on World Bank country level data on researchers per million. Model (5) is the base specification plus a control variable for overall immigration to the US based on country of origin (data sourced from the UN). Model (6) is the base specification including US State Dept Visas per country for work/exchange categories. Additional robustness checks were run and are discussed in the paper. All results in this table include firm and year fixed effects as well as robust clustered standard errors (clustered based on country).

Table 7: Influence of Subsidiary on the Ratio of Citations to Host Country over Citations to HQ (DV columns 1-3) and on the Ratio of Similarities (DV columns 4-6)

	(1)	(2)	(3)	(4)	(5)	(6)
Log(Mass Subsidiary)	0.201 *	0.204*	0.168*	0.470***	0.467***	0.373***
	(0.083)	(0.084)	(0.069)	(0.095)	(0.093)	(0.095)
Log(Mass HQ)		-0.072	-0.066		0.103	0.120
		(0.054)	(0.051)		(0.157)	(0.156)
Log(Country Mass)			0.045 *			0.117**
			(0.021)			(0.034)
_cons	-0.074	0.331	0.022	1.236 ***	0.664	-0.138
	(0.049)	(0.280)	(0.176)	(0.315)	(0.838)	(0.745)
Year FE	yes	yes	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes	yes	yes
N	3157	3157	3157	3157	3157	3157
R^2	0.067	0.068	0.075	0.053	0.0533	0.059
Clusters	69	69	69	69	69	69

Standard Errors in parentheses (country clusters)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

All models include firm and year fixed effects as well as robust clustered standard errors (country clusters). The first three models (1)-(3) use the ratio of subsidiary citations to the host country / (1+ subsidiary citations to HQ); the latter three models use the ratio of cosine similarities (similarity of subsidiary to host / similarity of subsidiary to HQ).

Table 8: Cosine Similarity of Cisco India to Host Country Compared to Average Similarity of Indian Subsidiaries to Host

	Cos Sim CISCO India to Host	Avg (Cos Sim Indian Sub to Host)
2005	0.2168	0.1177
2006	0.2255	0.1319
2007	0.3688	0.1800
2008	0.4686	0.2252
2009	0.4890	0.2634
2010	0.5783	0.3828
2011	0.5441	0.3631

Table 9: Cosine Similarity of IBM India to Host Country Compared to Average Similarity of Indian Subsidiaries to Host

	Cos Sim CISCO India to Host	Avg (Cos Sim Indian Sub to Host)
2005	0.3119	0.1177
2006	0.2985	0.1319
2007	0.4250	0.1800
2008	0.5030	0.2252
2009	0.6289	0.2634
2010	0.7116	0.3828
2011	0.8067	0.3631

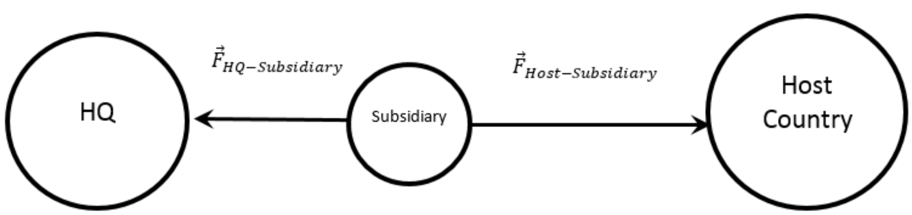

Figure 1. Schematic of two competing gravitational forces on the subsidiary.

APPENDIX

Multinational Firms in sample

Firm Name
IBM
MICROSOFT CORP
QUALCOMM
GOOGLE CORP
APPLE INC
GENERAL ELECTRIC
HEWLETT PACKARD
INTEL
GENERAL MOTORS
AT&T
BROADCOM
CISCO
MICRON
BOEING
XEROX
TEXAS INSTRUMENTS
FORD MOTOR
HONEYWELL
AMAZON
VERIZON
ORACLE
LSI
EMC CORP
FREESCALE SEMICONDUCTOR
3M