Proximity strategies in outsourcing relations: the role of geographical, cultural and relational proximity in the European automotive industry

*Alexander SCHMITT and
Jo VAN BIESEBROECK*

Econometrics

Faculty of Economics
And Business

Proximity strategies in outsourcing relations: The role of geographical, cultural and relational proximity in the European automotive industry[*]

Alexander Schmitt[1]
University of Leuven

Johannes Van Biesebroeck[2]
University of Leuven and CEPR

February 22, 2013

Abstract

Trends towards international fragmentation of production and modular process technologies have increased the importance of proximity in the supply chain of sophisticated manufactured goods. Using a rich and novel data set for the European automotive industry, we simultaneously evaluate the relative importance of geographical, cultural and relational proximity in sourcing strategies. The estimates indicate that each dimension provides an independent benefit and also which measures have the largest relative importance. We also find that the positive effects attributed to some measures reflect past relationships rather than predict new ones. In particular, co-location and a low cultural distance should be interpreted as outcomes of a sourcing strategy, not as predictors for sourcing success. We investigate to what extent firms from different countries follow different strategies and which choices suppliers can make to boost their attractiveness as outsourcing partner.

[1] Naamsestraat 69, 3000 Leuven, Belgium, tel. +32 16 373709, fax. +32 16 326796, E-mail: Alexander.Schmitt@kuleuven.be

[2] Naamsestraat 69, 3000 Leuven, Belgium, tel. +32 16 326793, fax. +32 16 326796, E-mail: Johannes.Vanbiesebroeck@kuleuven.be

1

1. INTRODUCTION

Sourcing strategies have changed fundamentally over the last decades. The production process of most goods has become increasingly fragmented in terms of ownership and global in reach and scope. The trade-in-tasks framework pioneered by Grossman and Rossi-Hansberg (2008) models this growing trend towards global sourcing of individual stages of production. Evidence collected by Johnson and Noguera (2012) shows a steadily falling share of value added in total trade flows. For many technologically advanced products, globalization of production has coincided with the emergence of lean and modular process technologies (Kotabe, Parente, and Murray, 2007; Sturgeon, Van Biesebroeck and Gereffi, 2008). This involves a more prominent role for external suppliers, who increasingly contribute to design and engineering. Suppliers are given more manufacturing and sub-assembly tasks and need to coordinate their production schedules with their clients.

In the North American and European automotive industries, outsourcing accelerated in the 1980s with the diffusion of Toyota's lean manufacturing techniques (Womack, Jones and Roos, 1990). Collaboration with outside suppliers increased both in development and production, which raised their relative importance.[3] In 2007, almost four fifths of the industry's employment in the United States was at supplier plants (Klier and Rubinstein, 2008). In Europe, several changes in the economic environment contributed to a reshaping of historical sourcing patterns. These include the integration of Central and Eastern European countries in the European Union, the establishment of local assembly plants by Asian firms, and a wave of mergers and acquisitions by the largest carmakers and suppliers. It makes the automotive industry a promising place to study how firms have adjusted to the aggregate trends mentioned above.

We estimate a model of supplier selection to learn how carmakers currently value proximity to suppliers. The work builds on the literature that measures the importance of bilateral distance in foreign market penetration (e.g. Berry, Guillén and Zhou, 2010), but we take the choice of supplier as the outcome of interest.[4] Our contribution is to apply a multi-dimensional concept of proximity. We analyze three dimensions—geographical, cultural and relational proximity—which are found to exhibit overlapping effects. Studying the importance of one proximity dimension without looking at the other two will lead to ambiguous conclusions and we illustrate several such instances. An additional contribution is to focus on several aspects of proximity that are under a supplier's control. The estimates reveal which proximity strategies are most effective for a supplier to boost its attractiveness as outsourcing partner.

Given the growing role of suppliers in the industry, it is natural to expect past relationships to confer current benefits. It is extremely rare, however, to observe systematic information on historical sourcing patterns. We use a novel measure of relational distance between suppliers and carmakers to show that proximity in this dimension is an important predictor of future contracts. One possible mechanism is the facilitation of knowledge exchange and the reduction of risk through trust (Dyer and Chu, 2000). Beyond the direct effect of relational proximity on sourcing, it is important to control for it when studying the importance of other determinants of supplier selection, such as geographical and cultural distance. Bönte

---

[3] Corswant and Fredriksson (2002) and Humphrey (2003) provide a detailed overview of the most important sourcing trends in the industry. The increased role of suppliers and the globalization of their activities are center stage.

[4] Our estimating equation can also be interpreted as a characterization of suppliers' market share in terms of proximity measures. The coefficient estimates are determined as the choice of suppliers by carmakers, and are, in turn, informative on the effectiveness of different supplier strategies.

(2008), for example, suggests a two-way interdependence between geographical proximity and trust in recurring relations.

The nature of automobile production has always required some geographical proximity and its importance is confirmed in our results. Transportation and logistics costs are naturally important for a final product that consists of thousands of parts. The growing role of suppliers in design and subassembly of components requires frequent interactions between suppliers and carmakers (Kotabe et al., 2007) and even among suppliers. Locating in a supplier park around an assembly plant is a popular strategy with multiple benefits (Frigant and Lung, 2002). In terms of sourcing success, however, we find that suppliers derive no independent benefits from co-locating with potential clients. It tends to reflect past relationships rather than predict future ones, as in the case of suppliers that follow their existing clients when they expand abroad (Martin, Mitchell, and Swaminathan, 1999). We also find that the benefits of locating in the proximity of complementary suppliers needs to be balanced by the business-stealing effects of suppliers that produce close substitutes.

Globalization and FDI have reshaped the organization of the European automotive industry and increased competition among parts suppliers. Market integration across the continent and the establishment of local assembly plants by foreign firms intensified contacts between firms with different cultural backgrounds. A preference for culturally close business partners has been a prevalent feature of the automotive industry, especially for Asian firms (Asanuma, 1989), driven among other things by a different perception of trust (Sako and Helper, 1998). Our estimates confirm that a shared nationality still provides suppliers with a strong benefit in sourcing relations. Relational proximity, as measured by the four dimensions in Hofstede (1980), is effective in overcoming cultural distance, but while it reduces the disadvantage of a different nationality, it does not eliminate it. Having an administrative presence in the assembly country, on the other hand, helps to overcome most of the remaining cultural gap.

We use a unique data source of individual outsourcing transactions that covers a large fraction of the European automotive industry. We assembled information on 19,323 contracts, each identifying the assembler of a particular car model, the component that is outsourced, and a list of potential suppliers for that specific component. The data set has a wide geographical reach, comprising 122 assembly plants and 1,530 supplier plants in 30 countries. In addition to the production locations, we collected information on the headquarter locations of all firms, assemblers and suppliers, as well as administrative branch offices of suppliers. Detailed ownership information links up all the elements in the data set.

Our model captures the extremely rich set of observed interactions in the sample. Multi-product carmakers own assembly plants in several countries and make repeated selections among a group of multinational component suppliers that operate production facilities in several countries and compete in several component markets. Studies of the geographical embeddedness of multinational firms have often lacked the analysis of strategic interactions among firms and between firms and locations (Beugelsdijk, McCann and Mudambi, 2010). Following the tradition in the economic geography literature, we construct a model of heterogeneous firms to measure interactions in space. We introduce additional explanatory factors drawing on insights from the international business literature. In particular, we highlight the simultaneous influence of spatial transaction costs and relationship assets in the strategic choices of firms. An advantage of the complex organizational structure of multinationals is that it allows us capture proximity effects along multiple dimensions and at multiple levels of detail: geographical proximity at the production or plant level, cultural proximity at the decision-making or headquarters level, and relational

proximity at the transaction or contract level. It improves the identification of causal effects (Corrado and Fingleton, 2012).

The remainder of the paper is organized as follows. In Section 2 we discuss why the automotive industry is a good place to study the impact of proximity on outsourcing relationships and why the 1993-2012 period in Europe is a good time and place to estimate our supplier choice model. This is followed, in Section 3, by the development of three testable hypotheses regarding proximity strategies. We provide information on the data set in Section 4 and discuss the empirical methodology and construction of bilateral proximity measures in Section 5. Estimation results follow in Section 6, and Section 7 concludes with a discussion of implications and limitations.

2. OUTSOURCING IN THE EUROPEAN AUTOMOTIVE INDUSTRY

Since the mid-1980s and throughout the 1990s, the automotive industry has undergone a fundamental restructuring in the way component production is organized. Traditional integrated production systems have given way to more dynamic disintegrated supply chains (Helper, 1991). Lead by Japanese car assemblers, so-called lean production systems are now implemented by car assemblers everywhere. Changes were first aimed at reducing in-process inventory costs with the development of just-in-time and just-in-sequence supply chain management systems. At the same time, production of more parts and of more sophisticated components was outsourced. Both trends greatly increased the need to coordinate with suppliers, not only in terms of logistics, but also in management and administration. Relationships with external suppliers, who took on more responsibility in the design and engineering process (Womack et al., 1990) deepened and expanded as supply networks evolved into sophisticated configurations. Choosing and maintaining the right supply base has become an important part of carmakers' strategy (Dyer, 1996) and proximity plays a crucial role in this.

The European automotive industry is particularly well suited to identify the effects of various dimensions of proximity on supply relations. It is one of the world's most competitive automotive markets with approximately 40 original equipment manufacturers (OEMs) and a large supply base (MacNeill and Chanaron, 2005).[5] Car assemblers produce multiple vehicles in assembly plants that are distributed across different countries. Production facilities of suppliers are also scattered across the continent to serve their various clients. In each component market it operates, a supplier will faces a different set of competitors which themselves vary in geographic and product scope. In this global industry, only a subset of the firms have regional headquarters in Europe and some suppliers have established local administrative branches to improve interactions with their clients.

Since the mid-1990s, the European automotive industry has undergone several changes that make the observed supply contracts to a large extent the result of actual choices and not merely legacy contracts with suppliers that were selected a long time ago (Lung, 2006). Three main trends deserve discussion.

First, the integration of new member states into the European Union (EU) coincided with rapid growth in their living standards and car sales. Local production was attractive for foreign multinationals to take

---

[5] OEMs are the organizational units of the car assemblers which design, produce, and market various car models under one brand (a single car assembler typically owns several OEMs). Since supply chains tend to evolve slowly over time and unique componentry is one way to differentiate brands, we consider the locations of the assembly plants as well as the (regional) headquarters of the OEMs in the analysis.

4

advantage of lower labor costs and the new plants could be used as export platform for the Western European market. This process started with Southern countries, Spain, Portugal, and Turkey, and accelerated when former communist countries in Central and Eastern Europe joined the EU, most notably Hungary, Slovakia, Slovenia, Romania, and Poland. Initially, new assembly plants were largely supplied from the OEMs' existing supply bases. As local component sectors developed, sourcing patterns on the entire continent adjusted to incorporate the best emerging suppliers (Bilbao-Ubillos and Camino-Beldarrain, 2008).

Second, following the 1992 EU single market program to integrate national economies more closely and the conclusion of the Uruguay round of multilateral trade negotiations in 1994, Asian firms established several new assembly plants in Europe. As a group, they operated only four assembly plants at the start of the sample period in 1993, but added ten new plants in subsequent years. This mirrored their earlier entry in the North American market. As they developed a local supply base, they had to decide for each component whether to rely on local firms or to pressure their existing suppliers to follow them to Europe. Increased price competition and falling trade barriers also induced car assemblers to increase component imports from emerging automotive sectors in large developing countries (Sturgeon and Van Biesebroeck, 2011).

Third, several multinational carmakers had to integrate newly acquired OEMs in their European production network. Ford established its Premier Automotive Group which included the formerly independent European firms Volvo, Jaguar, Aston Martin, and Land Rover. General Motors added Saab to its Opel/Vauxhall lineup and integrated the European operations of Daewoo with Chevrolet's. Volkswagen even combined eight OEMs in a single group. The integration of SEAT (from Spain) and Skoda (from the Czech Republic) enlarged its geographic scope, while the addition of four luxury and sport brands (Bentley, Bugatti, Lamborghini and Porsche) extended its presence beyond the mass-market segments. With each acquisition, an owner has to decide to what extent to integrate the supply chain of a newly acquired firm into its existing operations. To limit the number of distinct suppliers a firm needs to collaborate with, acquisitions tend to be accompanied by a rationalization of supplier relationships.

Each of these three trends prompted adjustments in European supply chains that affected all of the large component suppliers in our sample. We thus argue that modeling the configuration of observed outsourcing relationships as the outcome of a static supplier choice model is a reasonable approximation for the specific region and time period.

3. PROXIMITY STRATEGIES IN OUTSOURCING RELATIONS

3.1. The Importance of Supplier Proximity

A large number of studies investigate the make-or-buy decision of firms. The transaction cost economics literature in particular has generated many insights into factors that influence outsourcing or in-house production of parts and components, see Monteverde and Teece (1982) for an application to the automotive industry. When firms outsource production activities, they are certainly not indifferent which supplier to transact with. This selection depends on more than just prices and we investigate what suppliers can do to boost their attractiveness.

Car assemblers can choose from a competitive supply base with firms offering a broad range of components and support services. Important sources of competitive advantage include resource

endowments and efficiency (Wernerfelt, 1984) and the capability to meet standards of quality and reliability (Christopher, 2005). Throughout the analysis we control for these sources of competitive advantage that are constant across all a supplier's relationships. Similarly, we include country dummies to absorb factors that make some locations more attractive than others. Brakman, Garretsen, and van Marrewijk (2009) compare the relative importance of locations and linkages, while we merely acknowledging that fixed locational assets are important. Instead, we focus on the importance of another type of strategy that is novel in the study of outsourcing relations: the proximity between buyer and supplier.

Geographical, cultural and relational proximity can be considered as three separate strategic choices and their impact on supplier selection have each been studied separately. For example, Reichhart and Holweg (2008) discuss the important advantages that suppliers derive from locating in close proximity to their clients and even other suppliers. Sako and Helper (1998) argue that assembler's preference for suppliers from the same country stems from the greater facility of establishing trust between culturally close firms. Dyer (1996) illustrates the importance of relationships as firms that invest in building up specialized supply networks outperform competitors.

The relationship between geographic or cultural proximity and repeated interactions is likely to involve two-way causation which requires caution in interpreting empirical results (Bathelt and Glückler, 2003). In the absence of institutional or relational ties between firms initially, close proximity can be highly conducive to make a relationship work and provide a supplier with a competitive advantage. Such relationships that originate from convenience can lead to long-term collaboration, as former suppliers accumulate relationship-specific capital that transforms them to partners with unique advantages (Williamson, 1979). Physical proximity also helps mitigate lacking cultural or historical ties through its influence on the perceived psychic distance between firms (Håkanson and Ambos, 2010).

The influence of the different factors is likely to overlap as proximity exhibits multi-level, multi-dimensional effects. Cultural distance is more important at the decision-making level, but it tends to co-vary with geography which is more important at the production level. Relational proximity is likely to be most important at the contract level, but as it ties firms together over time, it leads to persistent effects of cultural and geographical advantages even after the underlying reasons for their importance might have disappeared. As a result, empirical evidence that measures the importance of one dimension but does not hold the other two dimensions constant will suffer from omitted variable bias.

The literature on foreign market penetration already emphasizes the multi-dimensional nature of proximity. The independent impact of several proximity dimensions on firm performance has received particular attention in the analysis of foreign consumer markets (Ghemawat, 2001) and FDI (Berry, Guillén and Zhou, 2010). Only by investigating the conditional importance of each dimension is it possible to identify the true importance of the underlying, independent forces. Otherwise, overlapping effects between different dimensions will lead to an upward bias and mitigating effects to a downward bias. Given the existing evidence obtained analyzing each dimension separately, the a priori expectation is for positive, independent effects for each of the three proximity dimensions.

Hypothesis 1: Geographical, cultural, and relational proximity each confer independent benefits that raise the attractiveness of a supplier as outsourcing partner, even after controlling for the other proximity dimensions.

3.2. The Independent Role of Geographical Proximity

The most straightforward reason why geographical proximity to a client is beneficial is that physical distance raises many transaction costs, such as transportation, logistics, and the costs or difficulties to meet delivery schedules (Duranton and Puga, 2004; Christopher, 2005). These costs are particularly important for components with a low value-to-weight ratio or for customized sub-assemblies that need to arrive at the assembly plant in a correct sequence (Womack et al., 1990). It is also easier for firms to search for suitable business partners among firms that are located nearby (Tabuchi, 1998). Given that all transaction costs raise the cost of the end product and are borne directly or indirectly by the outsourcing firm, they will influence its choice of supplier.

These effects are unlikely to be monotonic. Transportation costs are often concave with distance as the mode of transportation is adjusted for faraway shipments. In contrast, timely and frequent delivery schedules can becomes increasingly difficult to meet from distant locations. Another source of discontinuity are the costs associated with border crossings (Feenstra, 2002). They include logistical costs, inspections, and probability of delays, but also paperwork and regulatory compliance costs, which especially for foreign multinationals can be difficult to cope with.

Proximity also facilitates personal interaction and face-to-face communication which is required to inspect and monitor product quality, exchange tacit knowledge, and collaborate on problem solving. Specifically for the automotive industry, Dyer (1996) has shown that immediate physical proximity, as in the case of co-located supplier parks, leads to greater trust between contracting parties, among other benefits. Dyer and Chu (2000) argue that face-to-face communication is an important determinant of trust and Bönte (2008) finds direct evidence that physical proximity can induce inter-firm trust. While some of the communications-related benefits of clustering might have diminished in importance, the geographical organization of an increasingly complex, knowledge-based economy is still strongly affected by clusters (Porter, 2000).

Co-location of suppliers with assembly plants has emerged in the automotive industry to reap the efficiency and specialization benefits of outsourcing while still allowing close collaboration (Larsson, 2002; Sako, 2005). There are several confounding factors, however, that might explain the sourcing success of co-located suppliers, even in the absence of independent co-location benefits.

First, because supplier parks host several suppliers benefits could be due to improved interactions with the client or to an advantage of clustering with other suppliers. Both forces can confer different and independent benefits in a decentralized network structure, while in a centralized network structure clustering can reinforce the benefit of co-location (Lorenzen and Mudambi, 2012). Knowledge spillovers (Malmberg and Maskell, 2002) or productivity spillovers (López and Suedekum, 2009) have been shown to be important in manufacturing in general. Benefits specifically mentioned for automotive suppliers include synergies that facilitate design work, e.g. efficient exchange of information, access to high quality components, sharing of infrastructure, and facilitating technology adoption (Sako, 2005; Reichhart and Holweg, 2008). Many of these benefits could derive as much from proximity to other suppliers than to the client.

Second, it is important to recognize the simultaneous influence of agglomeration and dispersion forces in clusters (Ottaviano and Puga, 1998). As a predictor of outsourcing relations, the benefits of locating in a cluster need to be placed against the possible disadvantage of losing sales when locating close to competitors. This mirrors the business stealing effect that complicates the identification of productivity

spillovers from FDI (Altomonte and Pennings, 2009) or the difference between spillovers from horizontally and vertically related firms (López and Suedekum, 2009). Which type of outputs other suppliers produce in a cluster, in particular complementary or substitute parts, will influence the net benefits of a cluster.

Third, spatial clusters can also be viewed simply as a form of revealed preference for the properties of a location (Ellison and Glaeser, 1998; Brakman et al., 2009). Even in the absence of spillovers, a cluster variable will help control for unobservable attributes of a location that positively affect all firms located there and help isolate any independent benefits of co-location.

Fourth, when carmakers establish assembly plants in foreign markets, they often convince several of their existing suppliers to follow them (Martin et al. 1999). Those suppliers will initially be located nearby, not because of independent co-location benefits, but because they have not yet established other commercial ties in the new market. A related form of reverse causality is possible for leader firms, the first firm to establish an automotive assembly plant in a particular region. The first mover advantage allows a leader firm to establish a supply network that will serve its own needs best (Cantwell and Mudambi, 2011). This will not only provide a transportation cost advantage, but allows the firm to site its most trusted suppliers nearby and source knowledge-intensive parts from a local network. In contrast, laggard firm have to take the configuration of the supply chain in a region as given. It forces them to rely to a greater extent on their parent network, which by construction is located at greater distance.

In general, past or current sourcing contracts are likely to play a role when a supplier expands into a new location (Rosenbaum, 2012). Specialized firm-specific supply chain strategies can lead to co-location as a side effect, if they are part of a long-term strategy (Dyer, 1996; Kotabe et al., 2007). One thus needs to control for relational proximity before one can ascribe an independent causal effect of co-location on sourcing. Close proximity might result from the establishment of a relationship rather than causing it.

Hypothesis 2A: *Geographical proximity* to a client makes a supplier more attractive as outsourcing partner, but complete co-location does not provide independent benefits.

Hypothesis 2B: *Geographical proximity* to other suppliers confers benefits that make a supplier more attractive to clients, but the net advantage is diminished by the more intense competition in a cluster.

3.3. The Independent Role of Cultural Proximity

Some aspects of the proximity of two transacting parties transcend the location of production plants and are better captured at the firm level. One important dimension is the culture of firms. It includes practices and doctrines, social and economic goals, language, customs, the mentality that dominates among employees, and even the internal institutional environment and mode of interaction with the outside world. Benefits of a shared culture include direct savings in accounting and transaction costs and indirect benefits of aligning objectives and business practices. Because of the different aspects it embodies, from locally rooted cultural values to the psychic distance between firms' management, cultural proximity has generally been measured on multiple dimensions using survey data (e.g. Hofstede, 1980; Håkanson and Ambos, 2010).

Cultural distance is often included in the analysis with a transaction costs argument (Shenkar, 2001). Conducting business with culturally close partners lowers uncertainty and hence it lowers the likelihood

that agreements need to be revisited. By influencing the ease and incentive of acquiring and sharing tacit knowledge, it also raises asset specificity and helps sustain contracts. A shared culture facilitates collaboration and trust (Sako and Helper, 1998) and the establishment of a relational contract between two firms which is especially important when performance is multidimensional (Baker, Gibbons and Murphy, 2002). Culturally close suppliers require less monitoring, control, and transfers of competencies and skills. Many of these factors are particularly important in the automotive sector.

Following Kogut and Singh (1988), many studies in the international business literature have looked at the impact of cultural distance on multinationals' market entry or FDI decisions. We similarly expect cultural relatedness between firms to facilitate the governance and implementation of outsourcing contracts and to be a positive predictor of observed relationships. A shared nationality, as measured by the location of the firms' headquarters has already been shown to facilitate sourcing in the automotive sector (Asanuma, 1989). We consider two strategies for suppliers to boost their attractiveness to carmakers with whom they do not share a nationality.

First, suppliers can target clients that are culturally close. In the FDI literature it has already been shown that multinationals have more success in markets that are similar on several dimensions to their home market (Berry et al., 2010). There is reason to believe, however, that such proximity might not be as beneficial in the current context. Suppliers do not need to interact with a large number of consumers or local firms. They only need to connect and collaborate with one particular client firm, which itself is invariably a multinational enterprise that operates in many countries around the world (Sturgeon and Van Biesebroeck, 2011).

A second strategy for suppliers is to bridge the cultural gap rather than avoid it. They can establish an administrative branch in the neighborhood of the assembly plant to facilitate the resolution of design, production, and logistical problems which inevitably arise over the course of a supply contract. Before a relationship is initiated, supplier-carmaker interaction can be at the headquarter level, between the purchasing and sales departments, or between design and engineering divisions. Successfully fulfilling the contract, however, requires intensive collaboration with employees of the client at the assembly plant, which often is in a different country from the client's headquarter. Berry et al. (2010) and Ghemawat (1999) include administrative distance as important elements of proximity in their framework.

> **Hypothesis 3A**: *Cultural proximity* provides an independent benefit to a supplier that even a close relational proximity cannot fully overcome.

> **Hypothesis 3B**: Establishing an administrative presence in the country where a client operates is an effective strategy to bridge a cultural distance.

4. DATA

The information on outsourcing relations comes from *SupplierBusiness*, a consulting firm in the automotive industry. Each observation in the data set identifies a model-component-supplier triplet that represents an outsourcing contract between a car assembler and an external component supplier. The contract identifies the particular car or light truck model where the specified component will be installed. In total, we work with a sample of 235 models assembled in Europe between 1993 and 2012. We observe an average of 75 contracts per model, out of a total of 330 different components. The sample includes the 112 largest supplier firms active in Europe, which produce on average 31 different components and supply

on average seven models with each component. We have compiled information on all manufacturing locations, administrative branches and headquarters of suppliers at the city level. We also observe the assembly plants for all models and the carmakers' (regional) headquarters. We know for each model when production starts and when it is planned to end.

We now discuss each of the three elements in a contract triplet: models, components, and suppliers. Each car or light truck model carries the brand name of one of 40 OEMs, which in turn are owned by global parent firms. For example, Volkswagen AG is the largest group in our sample and is represented by five OEMs: Audi, Bentley, SEAT, Skoda and VW. We observe 42 distinct models for this firm and know in which of the 21 European assembly plants operated by Volkswagen AG each is assembled.[6] Information on headquarter locations of all OEMs comes from the *Automotive News* online data center. Table 1A lists the car firms in the sample, sorted by their total number of European assembly plants, along with the number of OEMs and distinct models. The last column shows the average number of unique components by model for which we observe the outsourcing contract.

[Insert Table 1 approximately here]

Components are classified into 330 unique categories, each belonging to one of the six major areas of a car: chassis, powertrain, exterior, interior, electronics, and miscellaneous. The categories take into account not only the generic component description, but also their general functionally and the area of the car where they are installed. We observe, for example, more than 600 contracts for bearings or bushes, but we create separate categories when they are installed in doors, wheels or engines, and whether they belong to the powertrain, chassis, or exterior areas.

Information on suppliers comes mostly from the *Amadeus* database compiled by Bureau van Dijk. It combines and harmonizes company accounts from all sectors of the economy and all European countries.[7] The supplier names on contracts are matched to firm names in several data fields of *Amadeus*. We learn the location of (regional) headquarters, of administrative branches listed in company reports, and of all production subsidiaries in 4-digit NACE codes associated with parts manufacturing for the automotive industry. Table 1B lists the largest supplier firms in our sample, which number 112 in total. The Schaeffler Group, for example, operates the most production facilities in Europe, 49, while the average is 14. It manufactures 54 distinct components and holds a supply contract for almost all of the car models in our data set (232). Unfortunately we do not observe which component is produced in which plant. To construct the distance between a model's assembly plant and the potential suppliers, we use the supplier's closest manufacturing facility.[8]

The sample contains a total of 122 assembly and 1,530 supplier plants. Table 2 summarizes their geographical distribution across Europe. Germany hosts the most assembly plants, one in five, followed by the United Kingdom and France. Fewer than one third of all assembly plants are located in the home

---

[6] A few models are assembled in more than one location, in which case we use the smallest average distance from these assembly plants to the suppliers' plants.

[7] Subject to a minimum size threshold, it provides comprehensive coverage of all firms that submit annual accounts to the national authorities. The threshold differs by country, but all firms that we are interested in easily exceed the minimum size threshold.

[8] One way to justify this approach is that plants can relatively easily produce several component and suppliers minimize transportation costs by allocating production to the closest location. The least demanding assumption that still allows consistent estimation is that the choice of specialization of suppliers' plants is uncorrelated.

country of the parent firm, mostly in the three countries with strong surviving OEMs: Germany, France, and Italy. Several plants are located in the country where the original OEM hails from, even though they are now owned by a foreign parent firm: United Kingdom, Sweden, the Czech Republic, and Romania.

The geographical distribution of supplier plants follows a similar pattern, but they are more dispersed and more likely to be foreign owned. While 21.6% of all European supplier plants are located in Germany, only slightly more than one third of them are owned by a German firm. The fraction of domestic ownership is even lower in all other countries. Outside Germany, only one supplier plant in twenty is located in its home country. To strengthen their corporate presence abroad, many suppliers have established administrative branches. Germany, the United Kingdom, and Spain are the most popular locations, but they are also prominent in countries with less developed local automotive sectors (the category *Other*).

[Insert Table 2 approximately here]

Figures 1A and 1B depict graphically the locations of, respectively, all assembly and supplier plants in the sample. They illustrate the extent of agglomeration. In Figure 1B we also indicate the presence of supplier clusters, which are described in the variables section. Comparing the two figures already hints at a close connection between the geographical distribution of assembly and supplier plants.

[Insert Figure 1 approximately here]

5. EMPIRICAL METHOD

5.1. Model

We estimate a supplier choice model where the dependent variable, a 0/1 dummy for sourcing success, has a supplier market share interpretation. The estimating equation describes the last stage in the carmaker's decision. A supply contract will specify a price, minimum quality, and delivery schedule. In addition, there is a more subtle, informal relational agreement (Baker et al., 2002), as the supplier is expected to collaborate with the client on design, innovation, and solve possible problems. The firm that is most likely to meet both the formal and relational terms, from the point of view of the carmaker, will be awarded the contract. We model this as an assembler choosing a supplier to maximize a latent expected profit function which is described by a reduced form equation that includes a rich set of proximity variables and controls.

Several choices have already been made prior to this. Suppliers have chosen the locations of production and administrative support offices as well as their product portfolio.[9] Carmakers have decided were to assemble each model and which components to outsource.[10] Their final decision is to select the preferred external supplier from the choice set, which we model using a conditional logit. The elasticities

---

[9] Rosenbaum (2012) incorporates the initial location choice of suppliers in a two-step estimation procedure, but still treats the product scope decision as predetermined, as we do.

[10] The make-or-buy decision is often modeled in a transactions cost framework, see for example Monteverde and Teece (1982). In the international trade literature the property rights theory is commonly used, see Antràs (2012) for a review and supporting evidence from trade in automotive products.

implied by the estimated coefficients are then used to make inferences about supplier strategy. Understanding what clients value is important for suppliers who wish to design a successful strategy.[11]

The explanatory variables of interest are the different dimensions of geographical, cultural, and relational proximity. Using fixed effects, we flexibly control for important mitigating factors, such as differences in supplier resources and capabilities, and constant locational advantages. These are important sources of comparative advantage that help a supplier on each contract. The conditioning in the model implicitly incorporates an intercept for each outsourcing contract, a model-component pair, thus holding the advantages of component-specific assets constant in the analysis. Details on the structural form of our model are the Appendix.

To estimate the conditional logit model, we need to define the choice sets of potential contracts from which a car assembler chooses one or more suppliers.[12] For each model-component contract, the choice set includes the supplier that is observed holding the contract, as well as all suppliers holding a contract to supply any of the 235 models in the data set with the exact same component. For each outsourcing contract there were on average 14 competing suppliers. The proximity variables of interest still vary across potential suppliers even after conditioning on each contract or choice set and including supplier fixed effects.

Figure 2 shows an excerpt of the data set to illustrate the key elements of the model. The first column enumerates the different choice sets, unique model-component pairs for which we observe a list of potential suppliers and over which the conditioning takes place (blue box). Each potential supplier for a contract enters as a separate observation. The dummy variable in the fifth column indicates whether a contract was signed or not—it is the outcome variable in the supplier choice model (green box). The last columns contain two examples of variables constructed to capture some dimension of bilateral proximity between a car assembler and potential supplier, the first at the plant and the second at the firm level (red box). The full set of these proximity variables is described in the next section.

[Insert Figure 2 approximately here]

We implement two estimators. First, we estimate the conditional logit model using maximum likelihood. It assumes that the error term that affects the competitive position of suppliers is identically and independently distributed across observations. This assumption is relaxed using a second, generalized method of moments (GMM) estimator.

The GMM estimator allows for a flexible, unobservable spatial pattern in the correlation across residuals. It allows us to separate the causal effect of the structural proximity variables from remaining spatial feedback mechanisms. McMillen (2012) argues that this is a good way to detect model

---

[11] A few studies have looked at the importance of distance in location decisions in the automotive industry. Smith and Florida (1994) explain the number of Japanese automotive-related manufacturing plants in US counties using the distance to the nearest Japanese assembly plants as explanatory variable. Klier (2005) uses the distance of each county to the city of Detroit as an explanatory variable for supplier employment. Klier and McMillen (2008b) model the location choice of automotive suppliers using a choice model with random alternatives and find that both the distance to Detroit and to the nearest assembly plant are important predictors. These studies use more aggregate models that cannot distinguish between the different dimensions we consider.

[12] In rare occasions, car assemblers decide to multiple-source a component, splitting one contract between two or more suppliers. This is allowed in the conditional logit model, but complicates the interpretation of the coefficients, see Richardson (1993) for a discussion.

misspecification. Corrado and Fingleton (2012) further highlight the importance of a hierarchical modeling of the spatial structure to improve the identification of causal effects. The inclusion of firm- and plant-level interaction variables in our estimating equation helps in this regard. In addition, the spatial filter applied to the model allows for spatial effects at even the most detailed (contract) level.

We build on the work of Kelejian and Prucha (1998) who propose the use of instrumental variables to recover a spatial autoregressive structure of first order. It involves first estimating the benchmark model (the conditional logit model in our case) and then using the fitted values to approximate the spatial term. Our implementation follows most closely the approximate GMM estimator of Klier and McMillen (2008a), as the size of our dataset does not permit an exact solution.[13] We adapted their estimator for the spatial logit model to the probability function of the conditional logit. It introduces an autoregressive term that captures the effect of a full-blown spatial structure. A statistical test on this term demonstrates that no residual spatial structure can be detected beyond the effects of our benchmark proximity variables.

In our application, a potentially more important source of correlation between the error terms of different observations stems from the existence of past relationships. A benefit of the GMM approach is that it can be generalized to other dimensions than geographical distance. In particular, we can model the error terms as having an autoregressive dependency on the intensity of past collaboration. The autoregressive term that we introduce in the preferred specification is defined as the average probability of sourcing success in any previous relationship between the current OEM and potential supplier.[14] It can be interpreted as the 'relational proximity' between the parties entering a particular contract. Details on the implementation are in Appendix A.3.

5.2. Variables

We characterize the proximity between car assemblers and potential suppliers using several bilateral distance and cluster variables defined for production plants, and three firm-level measures of culture. Variable definitions are summarized in Table 3.

[Insert Table 3 approximately here]

KM DISTANCE is calculated in a straight line between every assembly plant and the closest manufacturing plant of each potential supplier. Table 4 provides the sample average, standard deviation and correlations for the main variables. The mean distance is 367 kilometers, with a standard deviation of 667 km. It is highly skewed to the right as it takes on only positive values.

[Insert Table 4 approximately here]

From the distance variable, we generate two intuitive dummies that are motivated by the literature on the automotive industry. PROXIMITY 700KM and PROXIMITY 10KM indicate whether a supplier has a manufacturing plant within 700 or 10 kilometers of the assembly plant where the component must be

---

[13] An additional benefit of this approach is that it makes it possible to use the same outcome variable as in the maximum likelihood estimation. We had to adapt the estimator to the probability function of the conditional logit and provide details in the Appendix.

[14] Rather than having a weighting matrix that contains the inverse of geographical distances, we use the inverse of the average probabilities of selection in any of the earlier situations where an OEM could have selected a certain supplier. The statistical implementation is the same as in spatial case.

delivered. The first is a proxy for a single day's driving distance (Klier and Rubinstein, 2008); the second captures co-location of suppliers with the assembly plant in a supplier park, see Reichhart and Holweg (2008) for definitions and a typology. The large difference in cutoffs is to clearly distinguish between two underlying reasons for proximity: logistics costs versus the benefits of face-to-face interactions. The vast majority of potential suppliers locate within 700 km (86%); relatively few locate within 10 km (4.3%), but they still account for more than 11,000 observations. The NO BORDER variable indicates whether a supplier has a production plant in the country where the assembly plant is located. This is the case in just over half of the potential contracts (56%). These three variables allow for nonlinear effects of geographical proximity to the client.

Two variables capture the separate effects of locating in the proximity of horizontally related firms, i.e. other automotive component suppliers. SUPPLIER CLUSTER COMPLEMENTS indicates whether a plant is near at least two other suppliers that produce different components for different clients. The cutoff distance for a cluster is 10 km, consistent with the co-location variable. SUPPLIER CLUSTER SUBSTITUTES is constructed similarly, but counting only other suppliers that produce the same type of component.[15] On average, 45% of supplier plants are located within either kind of cluster, with complements the majority type (32%). Note the strong correlation between PROX 10KM and SUP CLUS COMP which is the result of supplier parks co-locating with an assembly plant. The inclusion of different cluster variables, in addition to the bilateral distance variables, allows us to capture a much more nuanced picture of the benefits of clustering. Recent research finds that multinationals exploit the advantages of each link in their larger geographical network (Jenkins and Tallman, 2010).

We also use three firm-level measures of proximity to capture factors that go beyond logistics and production spillovers. SAME NATIONALITY indicates whether the headquarters of the OEM and the supplier are in the same country. We use the nationality of the OEM rather than the parent firm because sourcing and supply chain decisions are to a large extend made independently by the OEM in charge of vehicle design. It means, for example, that Opel/Vauxhall, although owned by the U.S. firm General Motors, is classified as headquartered in Rüsselsheim, Germany. 15% of all potential suppliers share nationality with the OEM.

For OEM-supplier pairs with headquarters in different countries, we include a control for cultural proximity at the country level. HOFSTEDE MAHALANOBIS calculates the Mahalanobis distance over Hofstede's (1980) four cultural dimensions: uncertainty avoidance, power distance, masculinity, and individualism.[16] Following the critique of Shenkar (2001), recent research by Kandogan (2012) suggests that this aggregation improves on the metric by Kogut and Singh (1988). It takes into account the correlation pattern in the four dimensions. The variable is expressed in standard deviations relative to the sample mean and equals zero when OEM and supplier share nationality.

The third measure of cultural proximity is ADMINISTRATIVE BRANCH, which equals one if a supplier has an administrative branch or support office in the country where the model is assembled. They are only observed in foreign countries, i.e. outside of the country where a supplier's headquarters are located. One quarter of the potential contracts in the sample benefit from a branch.

---

[15] The results proved highly robust to alternative definitions; robustness checks are in the Appendix.

[16] The data was compiled by Prof. Boyd in "Hofstede's cultural attitudes research – cultural dimensions," http://www.boydassociates.net/Stonehill/Global/hofstede-plus.pdf, accessed March 23, 2012.

The last row of Table 4 shows the summary statistics for the dependent variable, CHOICE, an indicator whether a contract has been signed or not. Its construction has already been discussed in the methodology section. On average, about one in ten (9.1%) potential outsourcing contracts is actually signed.

6. RESULTS AND DISCUSSION

6.1. Main Results

Table 5 contains the implied semi-elasticities for the supplier choice model estimated using maximum likelihood. They indicate the effect of each proximity variable on the probability that an outsourcing contract is formed in percentage points.[17] In each specification, geographical proximity is a strong predictor of sourcing success. For example, specification (1) implies that locating 100km closer to an assembly plant is boosts the probability that a supplier is awarded the contract by 2.7%.[18] Specification (2) suggests that locating within a day's driving distance boosts the probability of sourcing success by 23%. The signs and magnitudes of these results are consistent with earlier studies.

[Insert Table 5 approximately here]

In line with the first hypothesis, we find that different proximity dimensions have overlapping effects. When variables are added in subsequent specifications, the point estimates change, generally in intuitive ways. For example, adding controls for supplier clusters reduces the estimated benefit of co-location with the assembly plant. In their absence, the coefficient on the PROX 10KM variable misleadingly reflected some of the positive spillovers from locating near complementary suppliers. When we introduce controls for cultural proximity, in specification (4), the estimated effect of co-location is reduced further.

The first column of Table 6 repeats the last specification of Table 5, but uses the approximate GMM estimator. The relative ranking of elasticities is virtually identical for both methods.[19] With this new estimator we can allow for unobserved spatial correlation in the residuals, which introduces a spatial autoregressive term in the equation. Its predictive effect on supplier selection is negligible and statistically insignificant, while the point estimates on the observed proximity variables barely change. It implies that they exhaustively characterize the effects of geographical proximity on supplier choice.

[Insert Table 6 approximately here]

As discussed earlier, we extended the spatial econometrics approach to control for correlation along a second dimension, namely the intensity of past collaborations between a supplier and OEM. This explicitly controls for the emergence of relationship-specific capital that can increase the likelihood of subsequent relationships for firms that have collaborated previously (Williamson, 1979). It also controls for unobservable factors that make a particular supplier-OEM match especially productive.[20]

---

[17] Where possible we use indicator variables to make it easier to compare absolute effects. Results including a continuous distance measure are described in the robustness checks in the Appendix.

[18] The number in parenthesis is the absolute value of the z-statistic of the corresponding point estimate. A z-statistic larger than 2, approximately, indicates that the average effect is statistically significant.

[19] The single instance where the relative magnitudes differ is statistically insignificant.

[20] It does reduce the sample size as we need to exclude the first observation for each supplier-client pair.

The estimates of our preferred specification, in column (4) of Table 6, provide strong support for the first hypothesis. Each of the three proximity dimensions has a strong, independent effect on sourcing outcomes. Locating within a day's driving distance (geography) and having headquarters in the same country (culture) are the two strongest predictors. Relational proximity, the last variable in the table, is also estimated to raise the likelihood of further sourcing success, even conditional on geographical and cultural proximity. The absolute magnitude of this last effect averages 7% over the full sample, but it can be as high as 54% for some contracts. The average importance of past relationships is approximately sufficient to overcome the disadvantage faced by a supplier with a different nationality but with a local administrative branch, or the disadvantage faced by a supplier located across a border.

The changes in coefficients when relational proximity is included underscore the importance of estimating all effects simultaneously. Because relational proximity is correlated with geographical and cultural proximity, the enduring effect of past relationships on current sourcing outcomes needs to be explicitly controlled for before one can assign a causal effect to those other factors. For example, Kumaraswamy, Mudambi, Saranga and Tripathy (2012) describe how local suppliers can leverage relationships with downstream multinationals to learn and catch-up with the technological frontier. Relationships provide access to advanced knowledge and access to the industry's global value chain, which strengthens a supplier's attractiveness to future clients. Nobeoka, Dyer and Madhok (2002) provide examples of suppliers that enlarge their client base with the explicit purpose of enhancing their technology learning. Co-locating production facilities or employees can play a crucial role in this strategy. Follow-sourcing, where suppliers follow existing clients overseas, is common in the automotive industry (Martin et al. 1999; Sturgeon and Van Biesebroeck, 2011).

The next two hypotheses concern the precise nature of the geographical and cultural effects. Suppliers that locate closer to a potential client gain benefits, but only up to a point. In particular, the positive effects of co-location on the sourcing probability should not be interpreted as an independent effect. It disappears almost completely when cultural and relational controls are included, which is something that previous studies have not done. Omitted variable bias is likely to be important in many settings as co-location will be correlated with a wide range of strategies. In contrast, the positive effect of locating within a day's drive remains remarkably stable when controls are added. The border effect declines in absolute importance once cultural proximity is controlled for, but it does retain a significant and independent effect throughout.

Once we distinguish between two types of supplier clusters, Hypothesis 2B also receives strong support. Proximity to firms producing parts with a different function in the final product—firms that are complementary in terms of technology and client-base—has a positive effect on sourcing outcomes, consistent with positive spillovers. Including this control diminishes the independent effect of co-locating with a client. Proximity to firms that produce substitute parts have a remarkably strong negative effect as business-stealing outweighs any possible spillovers. These two opposing effects of other suppliers can only be identified if the two dimensions of clusters are investigated separately, and if the effect of clustering near the client is simultaneously controlled for. Including only a single cluster variable, as is the case in most previous work, leads to much more muted average effects.

While locating 'close enough' is what mattered in the geographical dimension, being extremely close provides an additional advantage in the cultural dimension. Even though the SAME NATIONALITY variable is strongly correlated with relational proximity, and reduces in magnitude in the final specification, its independent effect on sourcing is still the largest of any included variable. The vast

majority of contracts is between two multinationals operating outside their home country. Only 5% of the observed relational ties are strong enough to generate a benefit of equal value as a shared culture.

A potential supplier with a nationality that differs from the client, but that is culturally similar (Hofstede, 1980) also has a higher probability of being awarded a contract. This mirrors the findings in Kirkman, Lowe, and Gibson (2006) who perform a meta-analysis evaluating the importance of this culture measure in a variety of settings. They find robust effects at the individual, organization, and country-levels of analysis, although culture explains only a small fraction of variation in most cases. In the Appendix, we show that the results are similar for an alternative measure based on the psychic distance approach of Håkanson and Ambos (2010). The evidence in our application, however, does not support the interpretation that cultural distance has an independent effect on sourcing. The effect is halved in magnitude and becomes statistically insignificant when relational proximity is controlled for. Like co-location, it mostly reflects a lingering effect of past relationships. It confirms the causal illusion that Shenkar (2001) warned for: a positive effect of low cultural distance should be considered an outcome of past relationships rather than a factor predicting them.

To understand the importance of proximity, it is also important to account for actions that firms can take to mitigate adverse effects. Given the importance of headquarter location, it is no surprise that suppliers with different nationalities often set up administrative branches in countries where they wish to supply assembly plants. The coefficient estimate on the ADMIN BRANCH variable suggests that such a strategy can almost entirely overcome the disadvantage of a different nationality. We have performed a similar analysis for administrative branches in the country where a potential client has its headquarters instead, but there the effect was much weaker. Consistent with Hypothesis 3B we find that support offices help, but they need rather be located close to a client's place of production, not its purchasing department.

In the Appendix we illustrate that these findings are robust to modifying the following three implicit assumptions of the model. First, given the apparent importance of proximity one might worry that suppliers choose locations or establish additional production facilities to increase the probability of being selected. Second, the definition of the set of unique components has implications for the substitutability of alternative suppliers and their presence in different choice sets. Third, the conditional logit model makes the independence of irrelevant alternatives assumption. By construction, the relative odds of choosing one supplier over another are invariant to the other suppliers in the choice set.

6.2. Heterogeneous Effects

We have seen that suppliers can adjust some of their proximity dimensions to make themselves more attractive outsourcing partners. The importance of some dimensions is likely to vary both across components and across potential clients. We now explore this heterogeneity.

Dyer (1996) has argued that inter-firm specialization is a source of competitive advantage. Many types of specific assets can lead to quasi-rents and enhance firm performance (Schoemaker and Amit, 1994). In our application, it is vital to distinguish between assets that are specialized to the component or to the local economic environment, and assets that originate from interactions between firms. The latter form the core of our proximity analysis, while the former are absorbed by model-component fixed effects

through conditioning.[21] Omitting these fixed effects is informative about any possible contract-specific heterogeneity in the importance of proximity dimensions.

One of the most important changes in column (3) of Table 6, relative to the preferred specification in column (4), is the reduced importance of the two geographical proximity variables: PROX 700KM and NO BORDER. For some components, alternative assets or capabilities are even more important than proximity and all potential suppliers tend to be located at a distance. A low wage cost or the availability of cheap energy are examples. If we pool observations from different contracts together as in column (3), the fact that some distant suppliers are successful leads to noticeably lower estimates on the geographic distance variables. Within a value chain, different types of linkages due to different technology or resource requirements have different proximity requirements that facilitate logistics and the transfer of knowledge (Fifarek and Veloso, 2010).

A similar mechanism reduces the coefficient estimate on relational proximity in column (3). For some basic parts with little quality variation, carmakers base their choice mainly on price and transportation costs. Other benefits such as the existence of unique assets pertinent to the component can be a good substitute for relational proximity. Sourcing will then shift towards the lowest bidder more frequently. As carmakers disregard past relationships, no close relationships emerge, which reflects in a lower estimate on the relational proximity variable.[22]

In both cases, failing to compare only within the relevant set of suppliers inappropriately treats too many firms as potential choices and biases the results. It confirms the experimental results in Buckley, Devinney and Louviere (2007). They study FDI location choices and show that it is crucial to consider the full set of available options to improve comparability in a world with heterogeneous preferences and boundedly rational managers. In our analysis, we find that to the extent that some proximity dimensions are of lesser importance in certain contracts and can be overcome by the benefits of specialized assets for the component, the estimated importance of those dimensions is biased downward if the supplier choice is not modeled over the relevant set of potential suppliers.

Different activities in global supply chains can contribute vastly different values to a vehicle, see for example Mudambi (2008) for knowledge-intensive products. We investigate how this influences the importance of proximity for supplier choice by introducing different coefficients by component type in the model. In column (9) of Table 7 we illustrate results for electronics components. Locating within a one day's drive and in the same country are the only factors that are (much) more important than in the benchmark specification. In contrast, a shared nationality or administrative branches are less important. This conforms well with the characterization of the modern electronics industry in Sturgeon (2002). Product design as well as manufacturing are technologically intensive and firms locate close to knowledge centers with highly skilled workers. Crucially, the codifiability of performance requirements and interfaces makes it easy to exchange knowledge over large distances and between firms with different cultures. Even though scale economies are large and one plant will serve many clients, the need for customization of automotive parts and integration in sub-assemblies makes it undesirable to locate

---

[21] Coefficient estimates from a conditional logit are identical to a logit model with contract fixed-effects, only the implied marginal effects will differ slightly (see Appendix).
[22] The increase in the point estimates of the PROX 10KM and the SAME NATIONALITY variables in column (4) is a mechanical result of the reduction in the estimated importance of relational proximity.

production facilities too far away. The optimal proximity strategy of suppliers clearly depends on the product lines they are active in.

Proximity effects might also vary by ownership if supply chain strategy differs for firms from different countries. Beugelsdijk (2007) has argued that it is important to allow for heterogeneity in firm strategy to reliably isolate the importance of (external) geography factors. The results in Table 7 allow for a more flexible model of the decision making process by interacting the proximity variables with the region of origin of OEMs or suppliers.

[Insert Table 7 approximately here]

The first three columns of Table 7 differentiate the elasticities by OEM, distinguishing between firms headquartered in Europe, North America and Asia. Given that the vast majority of potential contracts in the sample are with European OEMs, those elasticities are very similar to the benchmark results of Table 5. For North American OEMs, proximity within a day's drive is particularly important, but a shared nationality not. Both effects are exactly opposite for Asian OEMs, which are nearly 2.5 times more likely to award a contract to an Asian supplier. Their preference for co-located suppliers might again reflect endogenously chosen supplier locations. They also have the strongest preference for suppliers located within the same country, which could indicate that communication problems or foreign regulatory environments pose greater difficulties for them. The presence of an administrative branch only appears to be a factor when competing for the business of European OEMs, which is intuitive as it is their home market.

Differential cultural antecedents of supplier switching probability in turn influence where suppliers locate. Given that it is costly to reverse location decisions, suppliers will only choose to maximize the benefits of co-location if they anticipate long-lasting relationships. The importance of trust in supplier relationships of Japanese carmakers (Dyer and Chu, 2000; Sako and Helper, 1998) is more likely to induce some reverse causality and raise the coefficient estimate on the PROX 10KM variable. One should be cautious not to interpret the full effect as an independent benefit of co-location. In contrast, the 'close but adversarial' model of supplier relations practiced by American firms (Mudambi and Helper, 1998) should lead to lower estimated benefits of co-location, which is what we find. It does not necessarily imply that American carmakers place a lower independent value on co-location. It could simply be that suppliers are less likely to choose locations that maximize those benefits. This behavioral adjustment is consistent with the absence of any co-location effect for American suppliers in column (5).

The results distinguishing by supplier type, in columns (4) to (6), confirm that sharing a nationality with the OEM is very helpful for Asian, but not for North American suppliers. The much higher estimate for the co-location variable for Asian suppliers in column (6), compared to the estimate for Asian OEMs in column (3), suggests that co-location helps them even with OEMs of different nationality. Foreign administrative branches are especially effective for Asian suppliers, but even for North American suppliers they are the single most important proximity factor. Interestingly, when European suppliers set up branch offices in other European countries, this is perceived negatively by car assemblers and it lowers their probability of success. The importance of co-location or locating in the same country follows the same pattern as by OEM: it is vital for Asian suppliers, unimportant for American suppliers, and intermediate for Europeans. The importance of production facilities within one day's driving distance seems to originate mainly from contracts awarded to European suppliers.

19

We can also interact the proximity variables with a key feature of assembly plant location, whether it is in the home country of the parent firm or abroad. In the home country, the firm is more likely to be a leader, which is expected to facilitate establishment of tight connections with local suppliers and a local network of knowledge transfer. The distinction between leader and laggard firms manifest itself in two ways. On the one hand, geographical proximity to a client, especially co-location and lack of borders, is more important at home, in column (7), compared to abroad, in column (8). This is consistent with leader firms shaping local supply chains for own benefit. On the other hand, the advantage of co-location among suppliers is higher abroad than at home. The 'physical attraction' experienced by insiders that makes it difficult for outsiders to enter a network (Cantwell and Mudambi, 2011) is less of a barrier in supplier parks abroad. As the carmaker has less control over the local supply chain, it relies more on its own, more distant network, but also on the suppliers' network. The importance of culture also varies by location. Shared nationality and the efforts to bridge such gap (e.g. FDI, supplier foreign branches) are more important in the home country of the client, while cultural proximity plays a stronger role abroad.

7. IMPLICATIONS, LIMITATIONS AND CONCLUSIONS

Our analysis has highlighted the need to use a multi-dimensional concept of proximity when studying the determinants of outsourcing relationships. Carmakers value supplier proximity in geographical, cultural and relational space, but these different dimensions have overlapping effects. One should be cautious to attribute a causal effect to one dimension if the other dimensions cannot be controlled for, something often lacking in previous research. While the independent beneficial effect of having a production facility within a single day's drive of a client or within the same country is highly robust, this is not always the case. In particular, the observed relationship between sourcing success and co-location of suppliers and clients stems entirely from indirect effects of other variables. A shared nationality, the enduring influence of past relationships, and clusters of complementary parts suppliers each do have independent effects on sourcing outcomes. Co-location, on the other hand, is merely correlated with those underlying causes and should not be considered as providing an independent benefit in sourcing relations.

The findings also have practical relevance because many aspects of proximity are under a supplier's control. The estimated effects are informative about the relative effectiveness of different location strategies. We have found that a shared nationality has the single most important independent impact on sourcing success and suppliers can clearly not influence this. However, cultivating a close relationship with a carmaker substantially diminishes the disadvantage of a different nationality. This is particularly true working with Asian clients. The relationship-specific skills highlighted in Asanuma (1989) can to some extent be acquired. Establishing an administrative presence near an OEM's production site facilitates closer interactions and helps suppliers overcome most of the remaining cultural distance. American and Asian have used this strategy with great success to win contracts from European firms.

The results imply more generally that suppliers can pursue proximity strategies to tailor their attractiveness to different clients. These go beyond the cultivation of unique capabilities that can be deployed across all outsourcing relationships (Wernerfelt, 1984). An effective strategy needs to be tailored to the particular preferences of potential clients. Carmakers value specialized supply networks (Dyer, 1996) and elements of proximity can play a valuable role in this. Suppliers need to realize, though, that not all choices have a direct and independent influence on competitive advantage. Relational proximity, in

particular, confers positive, but only limited independent benefits. It can however leverage existing sources of comparative advantage that a supplier has in other dimensions.

The opening of new assembly plants in Eastern Europe induced carmakers to establish local supply networks. Past relations and cultural proximity was of crucial importance for suppliers entering these new locations, especially for system integrators (Bilbao-Ubillos and Camino-Beldarrain, 2008). Without relational proximity, establishing a presence in a co-located supplier park would save on transportation costs, but only confer small additional benefits through proximity to suppliers producing complementary parts. Cost savings from producing locally alone might not be enough to attract contracts from clients of a different nationality, lacking cultural proximity. Since OEM headquarters are firmly rooted in Western Europe, the periphery of the industry remains dependent on the core (Domanski and Lung, 2009). Suppliers that want to expand into emerging automotive industries need relational capital to leverage existing sources of comparative advantage and they should not neglect FDI in support offices to help overcome administrative distance.

The recent history of the North American automotive industry provides a dynamic illustration of the above effects. It was traditionally clustered around Detroit and the importance of geographical proximity seemed indisputable. The advent of modular production with a greater role for suppliers increased the relative importance of cultural and relational proximity, but given that most major suppliers were already headquartered in the Midwest, this organizational change initially had little impact on the industry's geography (Klier and Rubinstein, 2008). Since the 1980s, Asian and European carmakers established assembly plants in the South to avoid import quotas and take advantages of right-to-work laws. Initially, many suppliers supplied the new plants from more distant locations, the Midwest or even Japan. Culturally close suppliers were among the first to locate in the South (Smith and Florida, 1994). As the industry grew in the South, the most successful American suppliers built up relational capital and eventually moved production facilities there as well.

Our analysis also has implications for the ongoing debate on the role of distance in a globalized economy. In their bestsellers, both Cairncross (2001) and Friedman (2005) argue that although geographical and cultural proximity used to be important predictors of commercial relationships and trade flows, they are not so anymore. Reviewing the statistical evidence more systematically, Disdier and Head (2008) find a 'puzzling persistence' for the importance of distance in gravity equations, even in recent years and even for services trade. Our results suggest that the importance of past relationships for sourcing decisions can lead to enduring effects of proximity even if its independent benefit has diminished sharply. Compositional effects can also hide the reduced importance of distance for some products. Some highly complex products will always need proximity to facilitate collaboration (Leamer and Storper, 2001). For more basic products, however, our results are consistent with the availability of key assets or low production costs to trump the importance of proximity.

Our study has some limitations. First, our analysis implicitly separates the carmaker choices and the supplier strategies. Some decisions are likely to be taken jointly by both parties to a transaction. For example, the diminished role for co-location once relational proximity is taken into account could be explained by simultaneous decision-making. Addressing this issue ideally requires a fully dynamic model, which is beyond the scope of this study.

Second, we only evaluate the importance of proximity on the probability that a supplier can attract a new outsourcing contract, which can be interpreted as the impact on suppliers' market share. While this

dependent variable is a necessary condition for success, it does not capture all performance dimensions of interest, such as profitability, quality or the level of innovation. A more comprehensive analysis that includes alternative performance indicators and measures the complementarity between them is warranted. It would be interesting to revisit in our context Beugelsdijk (2007)'s finding of the secondary role of environmental variables in firms' innovation performance

Third, our sample only includes the largest suppliers. An important benefit is that these firms tend to a lot more clients (automotive assembly plants) than production locations. It limits reverse causality from sourcing contracts to locations as they can only co-locate with one of their many clients. It is possible, however, that proximity influences sourcing success of smaller firms in a systematically different way, for example because of the tiered organization of the supply chain in the automotive sector (Womack et al., 1990). We plan to investigate this issue in future research with access to better data.

REFERENCES

Altomonte, C. and E. Pennings (2009). Domestic plant productivity and incremental spillovers from Foreign Direct Investment. *Journal of International Business Studies* 40 (7), 1131-1148.

Antràs, P. (2012). Grossman-Hart (1986) goes global: Incomplete contracts, property rights, and the international organization of production. *Journal of Law, Economics and Organization*, forthcoming.

Asanuma, B. (1989). Manufacturer-supplier relationships in Japan and the concept of relation-specific skill. *Journal of the Japanese and International Economies* 3 (1), 1-30.

Baker, G., R. Gibbons, and K. Murphy (2002). Relational contracts and the theory of the firm. *Quarterly Journal of Economics* 117 (1), 39-84.

Bathelt, H. and J. Glückler (2003). Toward a relational economic geography. *Journal of Economic Geography* 3 (2), 117-144.

Berry, H., M. Guillén, and N. Zhou (2010). An institutional approach to cross-national distance. *Journal of International Business Studies* 41 (9), 1460-1480.

Beugelsdijk, S. (2007). The regional environment and a firm's innovative performance: A plea for a multilevel interactionist approach. *Economic Geography* 83 (2), 181-199.

Beugelsdijk, S., P. McCann, and R. Mudambi (2010). Introduction: Place, space and organization—economic geography and the multinational enterprise. *Journal of Economic Geography* 10 (4), 485-493.

Bilbao-Ubillos, J. and V. Camino-Beldarrain (2008). Proximity matters? European Union enlargement and relocation of activities. *Economic Development Quarterly* 22 (2), 149-166.

Bönte, W. (2008). Inter-firm trust in buyer-supplier relations: Are knowledge spillovers and geographical proximity relevant? *Journal of Economic Behavior & Organization* 67 (3-4), 855-870.

Brakman, S., H. Garretsen, and C. van Marrewijk (2009). Economic geography within and between European nations: The role of market potential and density across space and time. *Journal of Regional Science* 49 (4), 777-800.

Buckley, P. J., T. M. Devinney and J. J. Louviere (2007). Do managers behave the way theory suggests? A choice-theoretic examination of direct investment location decision-making. *Journal of International Business Studies* 38 (7), 1069-1094.

Cantwell, J. A., & Mudambi, R. (2011). Physical attraction and the geography of knowledge sourcing in multinational enterprises. *Global Strategy Journal*, 1 (3-4), 206-232.

Cairncross, F. (2001). *The death of distance: How the communications revolution is changing our lives*. Harvard Business Press.

Christopher, M. (2005). *Logistics and supply chain management: Creating value-added networks.* Harlow, UK: Pearson.

Corrado, L. and B. Fingleton (2012). Where is the economics in spatial econometrics. *Journal of Regional Science* 52 (2), 210-239.

Corswant, F. von and P. Fredriksson (2002). Sourcing trends in the car industry. A survey of manufacturers' and suppliers' strategies and relations. *International Journal of Operations & Production Management* 22 (7), 741-758.

Disdier, A. and K. Head (2008). The puzzling persistence of the distance effect on bilateral trade. *The Review of Economics and Statistics* 90 (1), 37-48.

Domanski, B. and Y. Lung (2009). The changing face of the European periphery in the automotive industry. *European Urban and Regional Studies* 16 (1), 5-10.

Duranton, G. and D. Puga. (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson and J.-F. Thisse (Eds.), *Handbook of regional and urban economics: Cities and Geography* 4, 2063-2117. Elsevier.

Dyer, J. (1996). Specialized supplier networks as a source of competitive advantage: Evidence from the Auto industry. *Strategic Management Journal* 17 (4), 271-291.

Dyer, J. and W. Chu (2000). The determinants of trust in supplier-automaker relationships in the U.S., Japan and Korea. *Journal of International Business Studies* 31 (2), 259-285.

Ellison, G. and E. L. Glaeser (1999). The geographic concentration of industry: Does natural advantage explain agglomeration? *The American Economic Review* 89 (2), 311-316.

Feenstra, R. C. (2002). Border effects and the gravity equation: Consistent methods for estimation. *Scottish Journal of Political Economy* 49 (5), 491-506.

Fifarek, B. J. and F. M. Veloso (2010). Offshoring and the global geography of innovation. *Journal of Economic Geography* 10 (4), 559-578.

Friedman, T. L. (2005). *The world is flat: A brief history of the 21st century*. New York: Farrar, Straus and Giroux.

Frigant, V. and Y. Lung (2002). Geographical proximity and supplying relationships in modular production. *International Journal of Urban and Regional Research* 26 (4), 742-755.

Ghemawat, P. (2001). Distance still matters. *Harvard Business Review* 79 (8), 137-147.

Grossman, G. and E. Rossi-Hansberg (2008). Trading tasks: A simple theory of offshoring. *American Economic Review* 98(5), 1978-1997.

Håkanson, L. and B. Ambos (2010). The antecedents of psychic distance. *Journal of International Management* 16 (3), 195-2010.

Hausman, J. and D. McFadden (1984). Specification tests for the multinomial logit model. *Econometrica* 52 (5), 1219-1240.

Helper, S. (1991). Strategy and irreversibility in supplier relations: The case of the U.S. automobile industry. *Business History Review* 56 (4), 781-824.

Hofstede, G. (1980). *Culture's Consequences: International differences in work-related values*. Sage, New York.

Humphrey, J. (2003). Globalization and supply chain networks: the auto industry in Brazil and India. *Global Networks* 3 (2), 121-141.

Jenkins, M. and S. Tallman (2010). The shifting geography of competitive advantage: clusters, networks and firms. *Journal of Economic Geography* 10 (4), 599-618.

Johnson, R. C. and G. Noguera (2012). Proximity and production fragmentation. *American Economic Review* 102 (3), 407-411.

Kandogan, Y. (2012). An improvement to Kogut and Singh measure of cultural distance considering the relationship among different dimensions of culture. *Research in International Business and Finance* 26, 196-203.

Kelejian, H. and I. Prucha (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics* 17 (1), 99-121.

Kirkman, B. L., K. B. Lowe, and C. B. Gibson (2006). A quarter century of "Culture's Consequences": A review of empirical research incorporating Hofstede's cultural values framework. *Journal of International Business Studies* 37 (3), 285-320.

Klier, T. (2005). Determinants of supplier plant location: Evidence from the auto industry. *Economic Perspectives* 3, 2-15.

Klier, T. and D. McMillen (2008a). Clustering of auto supplier plants in the United States. *Journal of Business and Economic Statistics* 26 (4), 460-471.

Klier, T. and D. McMillen (2008b). Evolving agglomeration in the U.S. auto supplier industry. *Journal of Regional Science* 48 (1), 245-267.

Klier, T. and J. M. Rubinstein (2008). *Who really made your car? Restructuring and geographic change in the auto industry.* W.E. Upjohn Institute for Employment Research.

Kogut, B. and H. Singh (1988). The effect of national culture on the choice of entry mode. *Journal of International Business Studies* 19 (3), 411-432.

Kotabe, M., R. Parente and J. Y. Murray (2007). Antecedents and outcomes of modular production in the Brazilian automobile industry: A grounded theory. *Journal of International Business Studies* 38 (1), 84-106.

Kumaraswamy, A., R. Mudambi, H. Saranga and A. Tripathy (2012). Catch-up strategies in the Indian auto components industry: Domestic firms' responses to market liberalization. *Journal of International Business Studies* 43 (4), 368-395.

Larsson, A. (2002). The development and regional significance of the automotive industry: Supplier parks in Western Europe. *International Journal of Urban and Regional Research* 26 (4), 767-784.

Leamer, E. and M. Storper (2001). The economic geography of the internet age. *Journal of International Business Studies* 32 (4), 641-665.

López, R. A. and J. Suedekum (2009). Vertical industry relations, spillovers and productivity: Evidence from Chilean plants. *Journal of Regional Science* 49(4), 721-747.

Lorenzen, M. and R. Mudambi (2012). Cluster, connectivity and catch-up: Bollywood and Bangalore in the Global economy. *Journal of Economic Geography*, forthcoming.

Lung, Y. (2006). The changing geography of the European automobile system. *International Journal of Automotive Technology and Management* 4 (2), 137-165.

MacNeill, S. and J. Chanaron (2005). Trends and drivers of change in the European automotive industry: Mapping the current situation. *International Journal of Automotive Technology and Management* 5 (1), 83-106.

Malmberg, A. and P. Maskell (2002). The elusive concept of localization economies: towards a knowledge-based theory of spatial clustering. *Environment and Planning A* 34 (3), 429-449.

Martin, X., W. Mitchell, and A. Swaminathan (1999). Organizational evolution in the interorganizational environment: incentives and constraints on international expansion strategy. *Administrative Science Quarterly* 43 (3), 566-601.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics*, 105–142. New York: Academic Press.

McMillen, D. P. (2012). Perspectives on spatial econometrics: Linear smoothing with structured models. *Journal of Regional Science* 52 (2), 192-209.

Mehta, C. and N. Patel (1995). Exact logistic regression: Theory and examples. *Statistics in Medicine* 14 (19), 2143-2160.

Monteverde, K. and D. Teece (1982). Supplier switching costs and vertical integration in the automobile industry. *Bell Journal of Economics* 13 (1), 206-213.

Mudambi, R. and S. Helper (1998). The 'close but adversarial' model of supplier relations in the U.S. auto industry. *Strategic Management Journal* 19 (8), 775-792.

Mudambi, R. (2008). Location, control and innovation in knowledge-intensive industries. *Journal of Economic Geography* 8 (5), 699-725.

Nobeoka, K., J. H. Dyer, and A. Madhok (2002). The influence of customer scope on supplier learning and performance in the Japanese automobile industry. *Journal of International Business Studies* 33 (4), 717-736.

Ottaviano, G. and D. Puga (1998). Agglomeration in the global economy: A survey of the 'New Economic Geography.' *The World Economy* 21 (6), 707-731.

Porter, M. (2000). Location, competition, and economic development: Local clusters in a global economy. *Economic Development Quarterly* 14 (1), 15-34.

Reichhart, A. and M. Holweg (2008). Co-located supplier clusters: forms, functions and theoretical perspectives. *International Journal of Operations & Production Management* 28 (1), 53-78.

Richardson, J. (1993). Parallel sourcing and supplier performance in the Japanese automobile industry. *Strategic Management Journal* 14 (5), 339-350.

Rosenbaum, T. (2012). Where do automotive suppliers locate and why? Yale University, *mimeo*.

Sako, M. (2005). Governing automotive supplier parks: Leveraging the benefits of outsourcing and co-location. University of Oxford, mimeo.

Sako, M. and S. Helper (1998). Determinants of trust in supplier relations: Evidence from the automotive industry in Japan and the United States. *Journal of Economic Behavior and Organization* 34 (3), 387-417.

Shenkar, O. (2001). Cultural distance revisited: Towards a more rigorous conceptualization and measurement of cultural differences. *Journal of International Business Studies* 32 (3), 519-535.

Smith, Jr., D. and R. Florida (1994). Agglomeration and industrial location: An econometric analysis of Japanese-affiliated manufacturing establishments in automotive-related industries. *Journal of Urban Economics* 36 (1), 23-41.

Schoemaker, P. and R. Amit (1994). Investment in strategic assets: Industry and firm level perspectives. In P. Shrivastava, A. Huff and J. Dutton (eds.), *Advances in Strategic Management*. JAI Press, Greenwich, Ct, Vol. 10A, 3-33.

Sturgeon, T. J. (2002). Modular production networks: a new American model of industrial organization. *Industrial and Corporate Change* 11 (3), 451-496.

Sturgeon, T. J. and J. Van Biesebroeck (2011). Global value chains in the automotive industry: An enhanced role for Developing Countries? *Journal of Technological Learning, Innovation and Development* 4 (1-3), 181-205.

Sturgeon, T. J., J. Van Biesebroeck, and G. Gereffi (2008). Value chains, networks and clusters: Reframing the global automotive industry. *Journal of Economic Geography* 8 (3), 297–321.

Tabuchi, T. (1998). Urban agglomeration and dispersion: A synthesis of Alonso and Krugman. *Journal of Urban Economics* 44 (3), 333-351.

Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal* 5 (2), 171-180.

Williamson, O. (1979). Transaction-cost economics: The governance of contractual relations. *Journal of Law & Economics* 22 (2), 233-261.

Womack, J. P., D. T. Jones and D. Roos (1990). *The machine that changed the world: How Japan's secret weapon in the global auto wars will revolutionize Western industry*. Harper Perennial.

Woodward, D. (1992). Locational determinants of Japanese manufacturing startups in the United States. *Southern Economic Journal* 58 (3), 690-708.

Table 1A. Car assembly firms in Europe

Number of	Assembly plants	OEMs	Models	Unique components per model [a]
VW	21	5	42	98
Ford	17	6	35	85
Fiat	15	5	34	60
Renault-Nissan	14	3	25	83
GM	12	3	17	76
PSA	10	2	25	90
Daimler	8	4	19	89
BMW	7	3	12	105
Toyota	7	1	8	67
Porsche	3	1	7	85
Chrysler	2	2	2	64
Suzuki	2	1	3	29
Honda	1	1	2	80
MG Rover	1	1	2	51
Tata	1	1	1	82
Hyundai	1	1	1	52
Total	**122**	**40**	**235**	**330**
Average	8	3	15	75

Note: [a] Average rounded to nearest integer.

Table 1B. Top supplier firms in Europe

Number of	Production plants	Unique components	Models per component [a]
Schaeffler	49	54	4
Continental	48	173	1
Bosch	47	96	2
Magneti Marelli	33	66	4
Total	32	86	3
Magna	29	107	2
TRW	26	109	2
Denso	22	67	3
Delphi	21	118	2
Valeo	21	94	2
ThyssenKrupp	21	87	3
Federal-Mogul	18	61	4
Visteon	13	75	3
Dana	13	64	4
Cooper Standard	6	74	3
(112 firms total)	…	…	…
Total	**1,530**	**330**	**1**
Average	14	31	7

Table 2. Geographical distribution of assembly plants, supplier plants and branches.

% by country	OEM assembly plants			Suppliers		
	All	Domestic OEM	Domestic parent firm	All plants	Plants in home country	Foreign administrative branches
Germany	20.5	15.6	13.9	21.6	8.9	16.3
United Kingdom	15.6	8.2	0.8	12.8	0.6	11.3
France	12.3	9.0	9.0	11.3	2.5	9.7
Italy	9.8	7.4	7.4	7.8	1.0	2.3
Spain	9.0	0.8		9.0	0.3	10.5
Turkey	4.9			1.8		
Czech Republic	3.3	1.6		5.5		0.8
Belgium	3.3			4.1		3.5
Hungary	3.3			2.8		
Poland	3.3			4.3		3.9
Portugal	3.3			2.0	0.1	1.2
Sweden	2.5	2.5		1.8	0.6	3.9
Romania	1.6	1.6		1.7		
Other	7.3			13.7	0.6	36.6
Total (%)	**100.0**	**46.7**	**31.1**	**100.0**	**14.6**	**100.0**
Total (No.)	122	57	38	1,530	224	257

Notes: Empty cells are zero entries. Other countries include Bulgaria, Croatia, Denmark, Finland, Greece, Ireland, Lichtenstein, Luxembourg, Malta, Morocco, Norway, Russia, Serbia, Slovenia, Switzerland, Tunisia and Ukraine.

Table 3. Summary of the variables

Variable	Type	Description
KM DISTANCE	Continuous	Assembly plant–supplier plant straight line distance, in kilometers.
PROXIMITY 700KM	Binary	Closest supplier plant is within 700km of assembly plant.
PROXIMITY 10KM	Binary	Closest supplier plant is within 10km of assembly plant.
NO BORDER	Binary	Supplier has a plant in country of assembly.
SUPPLIER CLUSTER COMPLEMENTS	Binary	Supplier plant is located within 10km of at least two other suppliers of different components (to different car assemblers).
SUPPLIER CLUSTER SUBSTITUTES	Binary	Supplier plant is located within 10km of at least two other suppliers of the same type of component (to different car assemblers).
SAME NATIONALITY	Binary	Supplier and OEM headquarters are in the same country.
HOFSTEDE MAHALANOBIS	Continuous	Mahalanobis distance over 4 dimensions of cultural proximity (Hofstede, 1980), standardized.
ADMINISTRATIVE BRANCH	Binary	Supplier has an administrative branch the country of assembly.
CHOICE	Binary	*Dependent variable*, supplier chosen to sign outsourcing contract or not.

Table 4. Sample correlations

Obs. 269,608	Mean	St. Dev.	DIST-ANCE	PROX 700	PROX 10	NO BORDER	CLUS COMP	CLUS SUBS	NAT-ION	HOF-STEDE	BRA-NCH
KM DISTANCE	366.7	666.7									
PROXIMITY 700KM	0.859	0.348	-0.583								
PROXIMITY 10KM	0.043	0.202	-0.115	0.085							
NO BORDER	0.565	0.496	-0.340	0.393	0.185						
SUP CLUS COMP	0.319	0.466	-0.111	0.125	0.167	0.167					
SUP CLUS SUBS	0.130	0.336	0.018	-0.024	0.004	-0.037	-0.264				
SAME NATIONALITY	0.153	0.360	-0.053	0.042	0.079	0.142	0.050	-0.012			
HOFSTEDE MAHAL	1.653	1.000	0.059	-0.097	-0.076	-0.217	-0.056	0.008	-0.697		
ADMIN BRANCH	0.253	0.435	-0.042	0.113	-0.021	0.214	0.057	-0.014	-0.169	-0.007	
CHOICE	0.091	0.288	-0.019	0.033	0.033	0.047	0.058	-0.079	0.113	-0.089	0.032

Table 5. Benchmark results: effect of proximity on the supplier choice probability

	(1)	(2)	(3)	(4)
DISTANCE (in 100km)	-0.0272			
	(10.16)			
PROXIMITY 700KM		0.227	0.230	0.232
		(6.436)	(6.457)	(6.455)
PROXIMITY 10KM		0.351	0.294	0.166
		(10.10)	(8.467)	(4.966)
NO BORDER		0.194	0.153	0.0469
		(9.058)	(7.208)	(2.278)
SUPPLIER CLUSTER COMPLEMENTS			0.231	0.195
			(13.33)	(11.32)
SUPPLIER CLUSTER SUBSTITUTES			-0.637	-0.639
			(30.98)	(31.09)
SAME NATIONALITY				0.567
				(12.42)
HOFSTEDE MAHALANOBIS				-0.123
				(7.454)
ADMINISTRATIVE BRANCH				0.201
				(6.673)
Pseudo-R² (fit)	0.0969	0.0983	0.114	0.126

Notes: 269,608 observations. The reported coefficients give the relative change in conditional logit probability with respect to a discrete change in the explanatory variable, evaluated at the sample average. Z-statistics (absolute value) in parenthesis. Supplier and country fixed effects included throughout. Contract-specific fixed effects absorbed through conditioning.

Table 6. Estimation of supplier choice probability controlling for correlation of residuals in geographic or relational space

	(1)	(2)	(3)	(4)
PROXIMITY 700KM	0.336	0.360	0.278	0.351
	(5.300)	(4.031)	(5.685)	(4.244)
PROXIMITY 10KM	0.123	0.111	0.089	0.064
	(3.475)	(2.347)	(2.120)	(1.432)
NO BORDER	0.100	0.099	0.0725	0.110
	(3.300)	(3.263)	(2.266)	(2.828)
SUPPLIER CLUSTER COMPLEMENTS	0.238	0.217	0.173	0.193
	(12.08)	(3.814)	(5.767)	(6.206)
SUPPLIER CLUSTER SUBSTITUTES	-1.576	-1.600	-1.120	-1.390
	(23.81)	(17.88)	(5.833)	(10.15)
SAME NATIONALITY	0.698	0.695	0.402	0.385
	(10.77)	(10.63)	(4.492)	(2.637)
HOFSTEDE MAHALANOBIS	-0.131	-0.132	-0.0850	-0.0751
	(4.043)	(4.049)	(2.681)	(1.481)
ADMIN BRANCH	0.373	0.378	0.193	0.303
	(8.074)	(7.958)	(3.410)	(4.817)
SPATIAL AUTOREGRESSIVE TERM:		0.000313		
(correlation of residuals in geographic space)		(0.394)		
RELATIONAL AUTOREGRESSIVE TERM:			0.313	0.560
(correlation of residuals in relational space)			(2.319)	(2.557)
Contract-specific fixed effects (conditioning)	Yes	Yes	No	Yes
Observations	269,608	269,608	186,506	184,930

Notes: The reported coefficients are approximate marginal effects on the conditional logit probability. T-statistics (absolute value) in parenthesis. Supplier and country fixed effects included throughout. Estimation with two-step GMM based on Klier and McMillen (2008a). P-value of Sargan test: 0.667 (1), 0.637 (2), 0.932 (3) and 0.617 (4). Further details on the estimation method in the Appendix.

Table 7. Effects interacted with different ownership, location and technology types.

	OEM headquarters			Supplier headquarters			Assembly location		Component category
	Europe	America	Asia	Europe	America	Asia	Home [a]	Abroad	Electronics
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PROX 700KM	0.251	0.372	-0.0996	0.311	0.0954	0.143	0.283	0.207	0.755
	(6.583)	(2.873)	(1.062)	(7.198)	(1.828)	(1.324)	(4.942)	(4.985)	(6.443)
PROX 10KM	0.155	0.113	0.196	0.200	0.0199	0.502	0.202	0.102	0.057
	(4.347)	(0.795)	(1.641)	(5.001)	(0.329)	(1.670)	(4.233)	(2.223)	(0.723)
NO BORDER	0.0514	-0.104	0.215	0.0585	-0.0197	0.235	0.0855	0.00899	0.163
	(2.285)	(1.699)	(2.695)	(2.371)	(0.617)	(2.598)	(2.693)	(0.316)	(3.204)
CLUS COMP	0.195	0.0321	0.176	0.221	0.133	-0.0256	0.152	0.224	0.163
	(10.60)	(0.556)	(3.019)	(10.03)	(4.626)	(0.315)	(6.558)	(9.397)	(4.026)
CLUS SUBS	-0.639	-0.734	-0.592	-0.670	-0.573	-0.727	-0.662	-0.622	-0.741
	(28.34)	(10.76)	(9.530)	(26.38)	(15.82)	(7.391)	(22.19)	(22.78)	(14.38)
SAME NATION	0.660	0.148	2.479	0.810	0.00062	2.315	1.006	0.334	0.259
	(11.17)	(1.200)	(9.413)	(13.38)	(0.0089)	(7.428)	(11.83)	(6.521)	(2.568)
HOFST MAHAL	-0.119	-0.179	0.0844	-0.0645	-0.170	0.0480	-0.0513	-0.147	-0.212
	(5.964)	(2.922)	(2.314)	(3.262)	(6.516)	(0.971)	(1.894)	(7.823)	(5.091)
ADM BRANCH	0.256	-0.120	0.0446	-0.276	0.277	0.374	0.319	0.132	-0.0914
	(7.825)	(1.627)	(0.650)	(3.024)	(7.641)	(3.868)	(8.387)	(3.684)	(1.852)
Subsample [b]	230,255	17,355	21,998	148,813	100,475	20,320	132,902	136,706	32,639

Notes: [a] Plant in same country as parent firm's headquarters, Europe only. [b] 269,608 observations per regression. Z-statistics (absolute value) in parenthesis. Supplier and country fixed effects included throughout. Contract-specific fixed effects absorbed through conditioning.

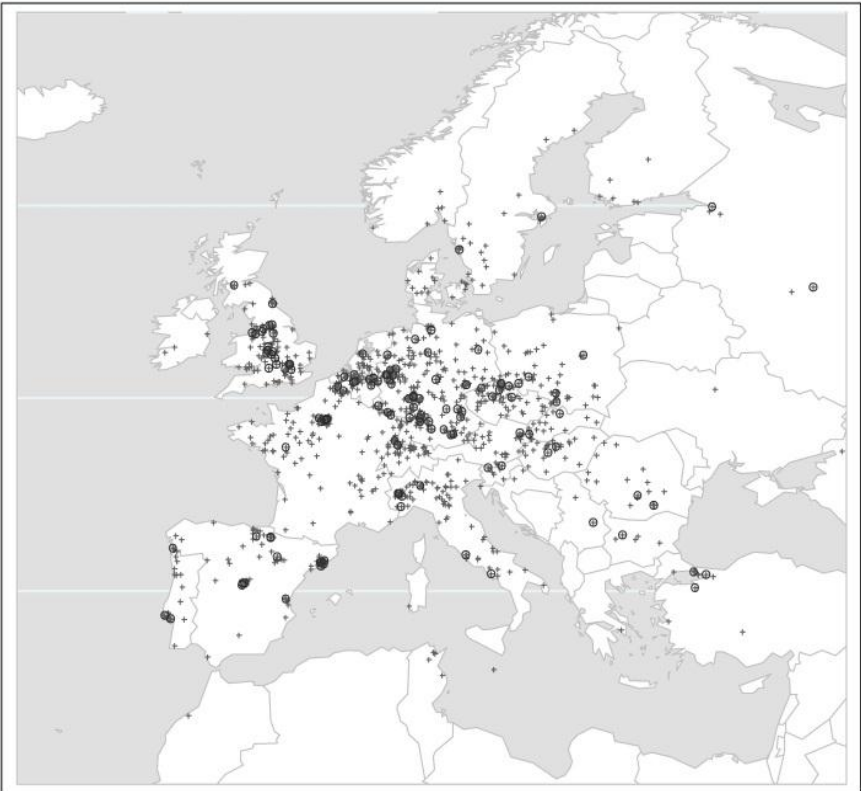**Figure 1A. Assembly plants.**

Figure 1B. Supplier plants and clusters.

Figure 2: Structure of the data set

Contract ID	OEM/Model	Component	Potential Supplier	Contract signed?	OEM/Model – Supplier Proximity Variables		
					NO BORDER	SAME NATIONAL.	Etc.
1	Ford Focus	Battery	Continental	0	1	1	...
1	Ford Focus	Battery	Delphi	0	1	0	...
1	Ford Focus	Battery	Johnson Cont.	1	1	0	...
2	Ford Focus	Tires	Bridgestone	1	0	0	...
2	Ford Focus	Tires	Continental	0	1	1	...
...	...	...	...	...	...	...	...
13	Audi A4	A/C pump	Behr	0	1	1	...
13	Audi A4	A/C pump	Delphi	0	1	0	...
13	Audi A4	A/C pump	Denso	1	1	0	...
13	Audi A4	A/C pump	Valeo	0	0	0	...
13	Audi A4	A/C pump	Visteon	0	0	0	...
14	Audi A4	Tires	Bridgestone	1	1	0	...
14	Audi A4	Tires	Continental	0	1	1	...
...	...	...	...	...	...	...	...

Notes: The sample contains a total of 269,608 observations (potential contracts), 19,323 unique outsourcing relations, 235 unique car models, 330 unique components, and 112 unique suppliers. Illustrative examples.

Unit of Analysis:
OEM/Model – Component

Key explanatory variables:
OEM/Model – Supplier proximity

Dependent variable:
Actual contract: YES/NO ?

A.1. Robustness Analysis

Measurements of geographic distance. We have conducted a series of robustness checks and report the relevant results in Table A.1. For all different specifications, a linear distance variable was included throughout. This tends to reduces the point estimates on the PROX700KM and NO BORDER variables, but leaves all qualitative findings unchanged—see for example specification (1).

We have also performed the regressions including an additional discrete distance effect, at 100km, which allows for multiple deliveries per day, or defined the cutoffs based on frequency distributions of distance rather than absolute distances. In all cases, the benefits of geographical proximity gradually decay as we move away from the assembly plant, at least if we do not control for relational proximity.

Alternative cluster variables. Next, we consider alternative measures of supplier clusters. In specification (2) we require a minimum of five (instead of three) plants to qualify as a cluster; in specification (3) we enlarge the radius around supplier plants that defines a cluster to 30km (from 10km). Both changes raise the number of firms in an average cluster, but in (2) the total number of clusters is reduced, while it is raised in (3). In both cases, the elasticities associated with any type of cluster increase, but the changes are very small. Using a more narrow definition of clusters leads to slightly lower elasticity estimates.

Definitions of cultural distance. We have also estimated the benchmark model using an alternative definition of bilateral headquarters' distance, grounded on the psychic distance approach. We have used the variable constructed by Håkanson and Ambos, defined as the "sum of factors (cultural or language differences, geographical distance, etc.) that affect the flow and interpretation of information to and from a foreign country" (2010: 201). For better comparison, we have standardized it relative to the sample mean. In specification (4) we see that this alternative variable produces similar results, but it captures less variation in our sample. Interestingly, it seems to capture more of the cultural aspects embedded in the border effect, but it performs worse at measuring very close cultures as in the case of shared nationality.

Correlation between contracts. The assumption that the residuals are uncorrelated might be violated if some important connection between observations is not modeled. The tests in Table 6 investigate this for two possible types of dependency, namely spatial correlation and historical linkages. Another potential violation of the independence assumption might occur if the supplier choice for one component depends on outcomes of contracts for other components of the same vehicle. We can accommodate such effects by using a clustered variance-covariance matrix that allows the covariance terms to differ from zero within each set of contracts for the same car model. The estimates in specification (5) of Table A.1 have indeed larger errors, but it does not affect the interpretation of the results.

Endogeneity of location. One potential endogeneity problem is that suppliers could change the location of their plants to influence the outcome of contracting decisions. It would invalidate our treatment of the observed locations as pre-determined supplier characteristics. Our analysis already mitigates this issue in two ways. First, by including contract-level fixed effects we only compare across suppliers for a given model-component pair which holds constant anything unusual about the contract, such as a remote assembly plant location, or a component with strong co-location requirements. Second, given that the sample includes for all suppliers many more contracts than production locations, they can locate their plants only near a few assembly plants.

Nonetheless, we can verify the sensitivity of the results by re-estimating the model excluding components that are more likely to suffer from this endogeneity problem. Components that are bulky, that

generate high coordination costs, or that require a lot of face-to-face interactions are candidates. In specifications (6) and (7), we re-estimate the benchmark specification, but exclude engineering-intensive or design-intensive components. Taking an extreme position, we eliminate half of the observations in the first regression and almost 40% in the second. The results change only slightly except for the PROX 10KM variable. It confirms our finding from Table 6 that endogenous locations cannot be dismissed entirely as an explanation for the effect of co-location.

Construction of choice sets. From the start, we had to define unique components to construct a set of potential suppliers for each observed contract. All results so far are based on a detailed classification system that separates components according to their generic name, functionality, and area of application in the car. An alternative approach is to group components according to their generic name alone. This broader definition groups components in the same category irrespective of their function or application in the vehicle. The implication is that choice sets include more potential suppliers for the same component. Competing suppliers will be less alike and contract-level fixed effects will absorb a smaller fraction of the variation. Results in column (8) of Table A.1 illustrate that using generic component names does not materially change the estimates.

The independence of irrelevant alternatives assumption. A well-known restriction of the conditional logit model is that it implies independence of irrelevant alternatives (IIA). By construction, the relative probability of selecting one supplier instead of another is independent of the presence or characteristics of further potential suppliers. Especially when the model needs to describe the choices of a heterogeneous group of decision makers over varying choice sets, this assumption might be overly restrictive. The inclusion of supplier-fixed effects alleviates this concern somewhat.

A statistical test for the validity of the IIA assumption is readily available (Hausman and McFadden, 1984). One needs to estimate the model excluding suppliers one by one from all choice sets, and then test for significant differences in the coefficient estimates. While excluding the majority of suppliers did not pose a problem, for a few of the largest suppliers in the sample and a few non-European firms the predictions in the benchmark model changed slightly but in a statistically significant way. Our specifications with interaction terms, however, provided more robust results.

Table A.1. Robustness analysis

	Bench.+ Distance	Alternative Variables	Cluster	Psychic Distance	Corr. Contract	Non Engineer	Non Designer	Generic Comp.
	(1)	(2)	(3)	(4)	(5) [d]	(6)	(7)	(8)
PROX 700KM	0.174 (4.481)	0.229 (6.410)	0.211 (5.931)	0.234 (6.330)	0.232 (4.592)	0.287 (5.404)	0.192 (4.261)	0.230 (6.351)
PROX 10KM	0.158 (4.719)	0.166 (4.950)	0.184 (5.492)	0.177 (5.217)	0.166 (3.099)	0.240 (5.129)	0.0987 (2.413)	0.164 (4.877)
NO BORDER	0.0332 (1.583)	0.0553 (2.682)	0.0515 (2.493)	0.0245 (1.135)	0.0469 (1.455)	0.128 (4.137)	0.0231 (0.919)	0.0703 (3.346)
CLUS COMPL	0.195 (11.36)	0.210 [a] (9.906)	0.241 [b] (13.07)	0.201 (11.32)	0.195 (7.661)	0.289 (10.59)	0.228 (10.26)	0.206 (11.71)
CLUS SUBST	-0.638 (31.06)	-0.529 [a] (19.19)	-0.555 [b] (31.41)	-0.638 (30.61)	-0.639 (23.15)	-0.557 (22.06)	-0.583 (26.98)	-0.587 (29.50)
SAME NATION	0.562 (12.34)	0.577 (12.65)	0.567 (12.41)	0.867 (19.95)	0.567 (6.815)	0.453 (7.256)	0.627 (10.75)	0.555 (12.04)
HOFST MAHAL	-0.123 (7.454)	-0.125 (7.629)	-0.120 (7.271)	-0.0395 [c] (2.389)	-0.123 (4.217)	-0.161 (7.012)	-0.0947 (4.477)	-0.125 (7.434)
ADM BRANCH	0.201 (6.681)	0.215 (7.125)	0.217 (7.123)	0.219 (7.003)	0.201 (4.409)	0.173 (4.335)	0.230 (5.788)	0.206 (6.861)
DISTANCE (in 100km)	-0.00807 (3.020)							
Pseudo-R² (fit)	0.126	0.116	0.130	0.125	0.126	0.159	0.142	0.116
Observations	269,608	269,608	269,608	255,479	269,608	137,257	166,231	290,117

Notes: [a] Min 5 cluster members. [b] 30km radius. [c] Håkanson and Ambos (2010: 202) metric, standardized. [d] Clustered errors (within uniquely defined car models). The reported coefficients give the relative change in conditional logit probability with respect to a discrete change in the explanatory variable, evaluated at the sample average. Z-statistics (absolute value) in parenthesis. Supplier and country fixed effects included throughout. Contract-specific fixed effects absorbed through conditioning.

A.2. Supplier Choice Model

For each set of potential contracts $r = 1,2,\ldots,n$ (defined in the methodology Section), car assemblers choose which suppliers $s = 1,2,\ldots,n_r$ they want to sign a contract with. We assume that the final choice of suppliers maximizes the profits of car assemblers. Denoting S_r the set of options available in the supplier selection process, the optimization problem is to choose which of the potential suppliers to award the outsourcing contract and which to decline it, so that expected profits are highest, given the quantities $\mathbf{q}$ specified in the contract.[23] Profits are modeled as a linear function of contract-specific characteristics α_r, which include the characteristics of car assemblers and assembly locations, a set of characteristics of suppliers and supplier locations $\boldsymbol{\beta}'\mathbf{x}_{rs}$, and a nuisance term ε_{rs} that captures unobserved factors plus measurement errors, assumed to be independently and identically distributed with a type-I generalized extreme value distribution:

$$\max_{S_r} \pi(S_r|\mathbf{q}) : \pi_{rs} = \alpha_r + \boldsymbol{\beta}'\mathbf{x}_{rs} + \varepsilon_{rs} \; \forall s \in S_r$$

McFadden (1974) shows how the maximization of a random utility function can be linked to the conditional logit model. Woodward (1992) applies an equivalent profit maximization problem to the location choice of Japanese manufacturing start-ups in the United States. Similarly, we study the geography of the European automotive industry with a model that implies an underlying profit maximization for car assemblers choosing amongst different component suppliers. The joint probability function that corresponds to such choices is given by:

$$\text{Prob}[Y_1 = y_1, \ldots, Y_N = y_N | m_1, \ldots, m_n] = \frac{\exp\left(\sum_{i=1}^{N}\sum_{s=1}^{n_r} y_{rs}(\alpha_r + \boldsymbol{\beta}'\mathbf{x}_{rs})\right)}{\sum_{r=1}^{n}\sum_{M_r} \exp\left(\sum_{r=1}^{n}\sum_{s=1}^{n_r} y_{rs}(\alpha_r + \boldsymbol{\beta}'\mathbf{x}_{rs})\right)}.$$

$Y_i = y_i$, $i_{\{rs\}} = 1,2,\ldots,N$ are the observed dichotomous yes-or-no choices of supplier, and m_r is the number of successful suppliers in each set of potential outsourcing contracts, over which the conditioning takes place. The two outer summations in the denominator are over the sets M_r of all possible combinations of 0 and 1 to $\{y_{r1}, \ldots, y_{rn_r}\}$ for all $r=1,2,\ldots,n$, such that $m_r = \sum_{s=1}^{n_r} y_{rs}$. Notice that the contract specific scalar α_r, as any other constant term in the model, cancels out after conditioning since it can be factorized away on both numerator and denominator. This implies that we are not able to retrieve its estimates. The parameter vector $\boldsymbol{\beta}$ is estimated by maximum likelihood estimation, as shown in Mehta and Patel (1995).

Our basic model specification includes K distinct measures of proximity and their average effects $\boldsymbol{\beta}' = (\beta_1, \ldots, \beta_K)$, $k=1,2,\ldots,K$, in addition to supplier fixed effects β_{9s} and country fixed effects for the locations of supplier plants β_{10c}:

---

[23] The underlying profit maximization problem of firm r is extended to incorporate supplier choice:

$$\max_{q,s \in S} \pi = p(q)q - C(z(q), w(z(q), s)),$$ with p and w for output and input prices, and q and z for output and input quantities; $C(\cdot)$ is the cost function.

$$\boldsymbol{\beta}'\mathbf{x}_{rs} = \beta_1 \text{PROX700}_{rs} + \beta_2 \text{PROX10}_{rs} + \beta_3 \text{NO BORDER}_{rs} + \beta_3 \text{CLUS COMP}_{rs} +$$
$$\beta_2 \text{CLUS SUBS}_{rs} + \beta_2 \text{NATION}_{rs} + \beta_2 \text{HOFSTEDE}_{rs} + \beta_2 \text{BRANCH}_{rs} + \Sigma_s \beta_{9s} + \Sigma_c \beta_{10c}$$

We report the mean of elasticity for each dichotomous proximity variable. More specifically, we calculate the transformed point estimates $g(\beta_k)$, where $g(\cdot)$ is a transformation that retrieves the change $\Delta y_i / y_i$ in the fitted outcome variable that is due to a change from *0* to *1* in a given variable x_{ki}.[24]

---

[24] We make the standard assumption of one positive outcome within each set of potential contracts, which is reasonable given that over 80% of the contracts in our sample are for single sourcing.

A.3. Two-Step GMM Conditional Logit Estimator

We can test whether we have omitted an extensive spatial correlation structure by checking if $\rho = 0$ in the following spatial autoregressive AR(1) model:

$$\mathbf{Y} = \rho \mathbf{W} \mathbf{Y} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}$$
$$= (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X} \boldsymbol{\beta} + (\mathbf{I} - \rho \mathbf{W})^{-1} \boldsymbol{\varepsilon}$$

In the second equation line, the inverse matrix has a full-blown MA(∞) representation that pre-multiplies both the error term and the explanatory variables:

$$(\mathbf{I} - \rho \mathbf{W})^{-1} = \mathbf{I} + \rho \mathbf{W} + \rho^2 \mathbf{W} \mathbf{W}' + \dots$$

We construct a spatial weight matrix using the inverse distance between the locations of suppliers, for each observation in the sample d_{ij}^{-1}, as shown below:

$$\mathbf{W} = \begin{pmatrix} 0 & d_{12}^{-1} & \cdots & d_{1n}^{-1} \\ d_{21}^{-1} & 0 & & \\ \vdots & & \ddots & \\ d_{n1}^{-1} & & & 0 \end{pmatrix}$$

If ρ significantly different than zero, the errors are not uncorrelated as assumed, and our βs will be inconsistently estimated due to omitted variables. While the first issue can be handled with bootstrap estimation, the problem of structural bias requires further treatment.

The variables $\mathbf{W} \mathbf{X}$ are easily obtained from the sample, but $\mathbf{W} \mathbf{Y}$ is not available, as $\mathbf{Y}$ is a latent variable. Kelejian and Prucha (1998) propose replacing $\mathbf{W} \mathbf{Y}$ by an instrumented variable. Klier and McMillen (2008a) demonstrate how a related two-step GMM estimator can be obtained for use with large samples, in order to avoid the need to work with very large matrices (for us this means inverting a square 269,608 matrix, which exceeds our computational resources). In their method, the objective function of the standard 2SLS-GMM estimator is replaced by a one shot guess of the orthogonality condition—a linear approximation at the point where $\rho = 0$.

The motivation is that the solution for the model $\mathbf{Y} = (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X} \boldsymbol{\beta} + (\mathbf{I} - \rho \mathbf{W})^{-1} \boldsymbol{\varepsilon}$ is relatively close to the point $\mathbf{Y} = \mathbf{X} \hat{\boldsymbol{\beta}}_0 + \hat{\boldsymbol{\varepsilon}}_0$. In the full spatial model, a generalized method of moments estimator would require to minimize the condition that the regressors $(\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X} \boldsymbol{\beta}$ are orthogonal to the error term $(\mathbf{I} - \rho \mathbf{W})^{-1} \boldsymbol{\varepsilon}$. In the linearized model, instead of using the full expression for the error term, it is approximated by applying a Taylor series expansion of first degree around $\hat{\boldsymbol{\varepsilon}}_0$:

$$(\mathbf{I} - \rho \mathbf{W})^{-1} \boldsymbol{\varepsilon} \approx \hat{\boldsymbol{\varepsilon}}_0 + \hat{\boldsymbol{\Gamma}}_0 (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_0)$$

Define $\mathbf{v} \equiv \hat{\boldsymbol{\varepsilon}}_0 + \hat{\boldsymbol{\Gamma}}_0 (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_0)$. Now the GMM method is to minimize the condition that a set of instruments $\mathbf{Z}$ is orthogonal to $\mathbf{v}$, with respect to $\boldsymbol{\theta}$:

$$\mathbf{v}' \mathbf{Z} \mathbf{M} \mathbf{Z}' \mathbf{v}$$

We use $\mathbf{M} = (\mathbf{Z}'\mathbf{Z})^{-1}$, which amounts to the two-stage least squares estimator. $\hat{\boldsymbol{\varepsilon}}_0$ and $\hat{\boldsymbol{\theta}}_0$ can be easily obtained from the initial conditional logit regression. $\hat{\boldsymbol{\Gamma}}_0$ is, by definition, the gradient of the error term of the full model with respect to the parameters of interest, an endogenous term if not treated. Measured at the starting point it becomes:

$$\hat{\boldsymbol{\Gamma}}_0 = \frac{\partial}{\partial \boldsymbol{\theta}} (\mathbf{I} - \rho \mathbf{W})^{-1} \boldsymbol{\varepsilon} \Big|_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_0}$$

Furthermore, the residuals in the full model $\mathbf{u} \equiv (\mathbf{I} - \rho \mathbf{W})^{-1} \boldsymbol{\varepsilon}$ can be retrieved by $\mathbf{u} = \mathbf{CHOICE} - \mathbf{P}$, where CHOICE is the observed binary choice variable. Thus the above reduces to:

$$\hat{\boldsymbol{\Gamma}}_0 = -\frac{\partial}{\partial \boldsymbol{\theta}} \mathbf{P}_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_0}$$

$\mathbf{P}$ is the conditional logit probability, again with the assumption of one positive outcome within each strata, written as, for each observation i:

$$P_i = \frac{\exp(\boldsymbol{\beta}' \mathbf{x}_i^{**})}{\sum^{n_r} \exp(\boldsymbol{\beta}' \mathbf{x}_i^{**})}, \text{ where } \mathbf{X}^{**} = (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X}^* \text{ and } \mathbf{x}_i^* = \sigma_i^{-1} \mathbf{x}_i, \; \mathbf{x}_i = (\mathbf{X}_i)'.$$

After some algebra, the gradient terms become:

$$\Gamma_{\beta_k i} = P_i \left[x_{ki}^{**} - \frac{\sum^{n_r} x_{ki}^{**} \exp(\boldsymbol{\beta}' \mathbf{x}_i^{**})}{\sum^{n_r} \exp(\boldsymbol{\beta}' \mathbf{x}_i^{**})} \right]$$

$$\Gamma_{\rho i} = P_i \left[\left(\boldsymbol{\beta}' \mathbf{x}_i^{***} - \sigma_i^{-2} \boldsymbol{\beta}' \mathbf{x}_i^{**} \operatorname{diag}(\boldsymbol{\Lambda})_i \right) - \frac{\sum^{n_r} \left(\boldsymbol{\beta}' \mathbf{x}_i^{***} - \sigma_i^{-2} \boldsymbol{\beta}' \mathbf{x}_i^{**} \operatorname{diag}(\boldsymbol{\Lambda})_i \right) \exp(\boldsymbol{\beta}' \mathbf{x}_i^{**})}{\sum^{n_r} \exp(\boldsymbol{\beta}' \mathbf{x}_i^{**})} \right], \text{ where:}$$

$$\mathbf{X}^{***} = (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{W} \mathbf{X}^{**}, \mathbf{X}^{**} = (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X}^*, \qquad \mathbf{x}_i^* = \sigma_i^{-1} \mathbf{x}_i, \qquad \mathbf{x}_i = (\mathbf{X}_i)' \text{ and}$$
$$\boldsymbol{\Lambda} = (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{W} (\mathbf{I} - \rho \mathbf{W})^{-1} (\mathbf{I} - \rho \mathbf{W})^{-1}.$$

Assuming at the starting point that $\rho = 0$, we do not have to bother about all the heteroskedastic terms, and the above gradients become much simpler, with $\boldsymbol{\Lambda}$ equal zero on the diagonal and $\mathbf{x}_i^{**} = \mathbf{x}_i$:

$$\hat{\Gamma}_{\beta_{0k} i} = \hat{P}_i \left[x_{ki} - \frac{\sum^{n_r} x_{ki} \exp(\boldsymbol{\beta}' \mathbf{x}_i)}{\sum^{n_r} \exp(\boldsymbol{\beta}' \mathbf{x}_i)} \right]_{\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}_0}$$

$$\hat{\Gamma}_{\rho_0 i} = \hat{P}_i \left[(\boldsymbol{\beta}' \mathbf{h}_i) - \frac{\sum^{n_r} (\boldsymbol{\beta}' \mathbf{h}_i) \exp(\boldsymbol{\beta}' \mathbf{x}_i)}{\sum^{n_r} \exp(\boldsymbol{\beta}' \mathbf{x}_i)} \right]_{\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}_0}, \text{ where } \mathbf{H} = \mathbf{W} \mathbf{X}, \; \mathbf{h}_i = (\mathbf{H}_i)'.$$

Now we have all the pieces needed to construct the variable $\mathbf{v}$ and run the 2SLS with instruments $\mathbf{Z}$. In the first stage, we regress the endogenous gradient terms on all our exogenous variables plus a set of

instrumental variables. In the second stage, the endogenous variables are replaced by the predicted values from the first stage.

Klier and McMillen (2008a) provide Monte Carlo results to show that this procedure (a logit variant) can deliver good results for parameter values $\rho < 0.5$. We estimate a ρ coefficient that is close to zero, positive and statistically insignificant. The instruments we used are all variables from the standard model, some of the **WX** variables that vary sufficiently in space, and the plain latitude and longitude coordinates. The validity of instruments is typically tested using over-identifying restriction in the GMM condition, a procedure called Sargan test. In the notes to Table 6 we list the p-values of this test, which does not reject the validity of our instruments.

The regression with the relational correlation variable uses essentially the same procedure, but instead of space, time is the dimension in the autoregressive matrix, called the lagging vector **L** now:

$$\mathbf{L} = \begin{pmatrix} L_1^1 \\ \vdots \\ L_n^1 \end{pmatrix}.$$

Because the time dimension in our sample does not exhibit sharply defined intervals, and because we want to estimate the broader effect of past collaboration, time is collapsed over all previous relationships between a certain OEM and a supplier into a single lagged period.

The estimates of the two-step GMM have the direct interpretation as marginal effects. However, the comparison with the results from the conditional logit model is not direct. The two-step GMM is an approximation method, with the variance-covariance matrix computed numerically, and well performing point estimates only if the sample is large enough (which is the case in our study). Still, the relative magnitude of the point estimates is very much comparable, as discussed in the results section.