Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Supply Chain Proximity and Product Quality

Robert L. Bray, Juan Camilo Serpa, Ahmet Colak

To cite this article:
Robert L. Bray, Juan Camilo Serpa, Ahmet Colak (2019) Supply Chain Proximity and Product Quality. Management Science 65(9):4079-4099. https://doi.org/10.1287/mnsc.2018.3161

informs
http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Vol. 65, No. 9, September 2019, pp. 4079–4099
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Supply Chain Proximity and Product Quality

Robert L. Bray,[a] **Juan Camilo Serpa,**[b] **Ahmet Colak**[c]

[a] Kellogg School of Management, Northwestern University, Evanston, Illinois 60208; [b] Desautels Faculty of Management, McGill University, Montreal, Quebec H3A 1G5, Canada; [c] College of Business, Clemson University, Clemson, South Carolina 29634

Contact: r-bray@kellogg.northwestern.edu, http://orcid.org/0000-0003-2773-0663 (RLB); juan.serpa@mcgill.ca, http://orcid.org/0000-0001-9737-2852 (JCS); acolak@clemson.edu (AC)

Received: September 15, 2017
Revised: April 12, 2018
Accepted: June 9, 2018
Published Online in Articles in Advance: May 17, 2019

https://doi.org/10.1287/mnsc.2018.3161

Abstract. We estimate the effect of supply chain proximity on product quality. Merging four automotive data sets, we create a supply chain sample that reports the failure rate of 27,807 auto components, the location of 529 upstream component factories, and the location of 275 downstream assembly plants. We find that defect rates are higher when upstream and downstream factories are farther apart. Specifically, we estimate that increasing the distance between an upstream component factory and a downstream assembly plant by an order of magnitude increases the component's expected defect rate by 3.9%. We find that quality improves more slowly across geographically dispersed supply chains. We also find that supply chain distance is more detrimental to quality when automakers produce early-generation models or high-end products, when they buy components with more complex configurations, or when they source from suppliers who invest relatively little in research and development.

History: Accepted by Vishal Gaur, operations management.
Funding: Financial support was received from Fonds de Recherche du Québec-Société et Culture [NP-207378].
Supplemental Material: The online appendix is available at https://doi.org/10.1287/mnsc.2018.3161.

Keywords: quality management • empirical operations • supply chain management • geographic proximity

1. Introduction

Proximity of upstream and downstream production plants should improve product quality because it facilitates communication, oversight, and responsiveness. However, studying the relationship between supply chain proximity and product quality is difficult, because doing so requires data on (1) a product quality metric, (2) the location of upstream factories, (3) the location of downstream assembly plants, and (4) product-level links connecting the two factories. To obtain these variables, we combine four independent automotive data sets that—despite their disparate origins—harmonize at the product level, yielding a panel of 28,500 supply chains:

- We use the *product quality metric* from (1) NHTSA's SaferCar data set, which reports 976,000 component defects, and (2) IHS's PolkInsight data set, which reports annual vehicle registrations. We measure component quality with defects per vehicle-year. For instance, the 2002 Ford Fusion's engine had 685 defect reports in 4,305,340 vehicle-years of drive time.
- We use the *upstream factory locations* from Bureau van Dijk's Orbis database, which reports the whereabouts of 26,375 facilities of 14,798 component suppliers. For example, Valeo Sylvania—a supplier of the exterior lighting system has factories in Seymour, Indiana, USA (latitude 38.94°N, longitude –85.89°W) and Queretaro, Mexico (–20.58°N, –100.38°W).
- We use the *downstream assembly plant locations* from SupplierBusiness's Who Supplies Whom (WSW) database, which reports the coordinates of 275 assembly plants. For example, Chevrolet has 25 assembly plants, including one in Gravatai, Brazil (latitude –29.94°N, longitude –50.99°W) and another in Kansas City, Kansas, USA (39.11°N, –94.62°W).
- We use the *product-level links* from WSW, which connects upstream suppliers with downstream automakers for 171,000 auto components. For example, Toyota installed 97 parts from Midway Products in the 2011 Corolla, including battery trays, child tether brackets, spring brackets, and rear stabilizers.

We merge these data sets to create a panel of 28,500 supply chains. For example, we observe that Infasco Nut produced fasteners for the 2014 Corvette in Mississauga, Ontario, Canada (latitude 43.74°N, longitude –79.64°W); Chevrolet installed the part at their Bowling Green, Kentucky U.S. plant (latitude 36.95°N, longitude –86.42°W); and customers filed 53 defect reports pertaining to this part over 1.363 million vehicle-years on the road.

The supplier-selection process makes the distance to the upstream factory endogenous—that is, both proximity and quality influence whether the supplier wins the procurement contract. We control for this sample selection bias in two ways. First, we use Dahl's (2002) selection model, instrumenting for the quality of the selected supplier with the characteristics of its competitors. Second, we treat plant relocations as quasi-exogenous supply chain distance shocks. In our sample, production of 79 car models moved from one assembly plant to another. These uprootings shifted the supply chain distances of the models' 2,451 components (53% of the supply chains lengthened; 47% shortened). We regress the change in the defect rates on the change in the supply chain distances pre- and postrelocation.

Both methods yield the same result: supply chain distance lowers product quality. Replacing the current supply chain distance with the distance of the next-farthest supplier—a 1,271-km increase, on average—increases the expected defect rate by 1.57%–2.17%. Tracking the evolution of component defect rates across production years, we find that quality improves more slowly when the supply chain is longer.

Finally, we study the moderators between distance and quality. We find that distance is more detrimental when the vehicle is early generation or high end, when the component is complex, or when the supplier is located is in another country or invests little in research and development. Thus, proximity is more salient for the electrical system of the 2010 Lexus CT—a complex component of a first-generation, luxury model—than for the interior lights of the 2011 Camry—a simple component for a seventh-generation standard model.

2. Literature Review

2.1. Supply Chain Management

Udenio et al. (2014) create a structural econometric model of the buyer–supplier links reported in Compustat's customer segment data; their model suggests that upstream inventory targets evolve dynamically with economic and financial conditions and systematically overreact to downstream inventory adjustments. With the same Compustat data, Serpa and Krishnan (2017) study efficiency spillovers across firms in the supply chain; they argue that productivity shocks propagate upstream from customer to supplier. Hendricks and Singhal (2005) also use Compustat data to show that upstream firms suffer from abnormal returns as a result of upstream disruptions. Tunca and Zhu (2018) study inventory financing in the supply chain spanned by online retailer JD.com and 170 of its suppliers; they show the retailer secures the smaller suppliers' loans and "demonstrate that [this] buyer intermediation induces lower interest rates and wholesale prices, increases order quantities, and boosts supplier borrowing." Jain et al. (2014, p. 1202) create a sample of buyer–seller supply chain dyads from 25.8 million U.S. Customs bill-of-lading manifests to study the impact of global sourcing on inventory investments. The authors estimate that "an increase in global sourcing results in an increase in inventory investment. A 10% shift in sourcing from domestic to global suppliers increases the inventory investment by 8.8% for an average firm." Wu (2016) combines these U.S. Customs dyads with supply chain links gleaned from Compustat and Bloomberg to construct a sample of one million buyer-supplier pairs; he also searches through news sources and firm disclosures to identify 8,000 firm-specific production shocks. Wu (2016, p. 2) estimates that production disruptions "cause substantial impact to firms even up to 4 connections away from the origin." Wang et al. (2017, p. 2) also study production shocks with Bloomberg's supply chain data. The authors show that downstream firms incur procurement risks when their suppliers' suppliers are concentrated. They find that a 10% increase in the commonality of tier-2 suppliers leads to a 4% increase in tier-0 supply risk. Osadchiy et al. (2016) use the Bloomberg supply chain data to study how different inventory echelons respond to business cycle fluctuations; they show that macro-level shocks have a larger effect on upstream suppliers. They attribute this result to order aggregation across firms and time.

Cachon and Olivares (2010) use *Automotive News* to link upstream automakers with downstream dealers, identifying two supply chain features that account for nearly all the cross-sectional variation in inventory holdings; they report that "the number of dealerships in a manufacturer's distribution network and a manufacturer's production flexibility explain essentially all of the difference in finished-goods inventory between Toyota and . . . Chrysler, Ford, and General Motors." Schmitt and Van Biesebroeck (2013, p. 202) use WSW and Orbis data to demonstrate the importance of proximity in auto supply chains. Rosenbaum (2014, p. 1) combines the WSW sample with Dun & Bradstreet data to estimate the trade-off suppliers face between labor costs and distance to the automaker. He estimates that "for a foreign-owned supplier, having a plant in a right-to-work state is equivalent to having a plant 2,000 miles closer to the assembler." Adams (2013) also uses Dun & Bradstreet data to structurally estimate the suppliers' facility location problem.

2.2. Product Quality

Like us, Ramdas and Randall (2008), Shah et al. (2016), and Colak and Bray (2016) study auto defect rates with NHTSA data. Ramdas and Randall (2008, p. 922) estimate the factors that influence Ford brake rotor defect rates. They find that the hazard rate of failure decreases with production volume and increases with the number of car models that use the part. Thus, they conclude that "the popular design strategy of component sharing can in some cases compromise product quality via reduced reliability." Shah et al. (2016, p. 3) study the operational drivers of recalls. Combining NHTSA's recalls data with Harbor Report's production data, the authors show that vehicle recall rates increase with factory utilization and the number of factory-installed options (such as moon roofs and dynamic steering). Specifically, they estimate that "a one-standard-deviation increase in the number of options (four additional options) is associated with two additional recalls." Colak and Bray (2016) structurally estimate the process by which customer failure reports translate into automaker and federal government recalls.

Guajardo et al. (2016) also study auto defect rates, but with a different data set: JD Power's Initial Quality Study. The authors find that warranties offset the effect of poor quality: they estimate that the median year-long warranty is worth roughly 3.1% of the median vehicle price. Guajardo et al. (2012) also study after-sales maintenance support; with a sample of aircraft engine maintenance contracts, the authors reveal that performance-based contracts (which peg service provider compensation to engine utilization) yield better performance than time-and-materials contracts (which peg service provider compensation to repair costs), thereby proving that service providers exert more effort under performance-based contracts.

3. Theoretical Motivation

Automakers prefer nearby suppliers: Boston Consulting Group (Spindelndreier et al. 2015) reports that proximity is the second-most-important auto part sourcing factor, and Berking et al. (2016) report that proximity is the third-most-important facility location factor.

This preference for proximity should in part arise from product quality considerations. First, it's easier to monitor nearby suppliers. "Co-location (i.e., the proximity of suppliers' operations to the automaker's), not only lowers monitoring costs," Gawer (2011, p. 255) argues, "but also facilitates the development of local norms" and "[p]roximity to plants makes it easier for headquarters to monitor and acquire information about plants," Giroud (2013, p. 861) finds. Salerno et al. (1998, p. 62) explain that "[f]rom the supplier's point of view, proximity represents a chance to improve the relationship with its client, which may lead to an adaptation of the component (or subset) design in order to solve some practical problems, or even co-design when developing new components for the local conditions." And Schmitt and Van Biesebroeck (2013, p. 479) mention that "proximity can be highly conducive to making a relationship work, and can provide a supplier with a competitive advantage. Such relationships that originate from convenience can lead to long-term collaboration."

Second, proximity leverages local market knowledge, as firms "located in proximity to other markets . . . will search more and find better suppliers" (Bernard et al. 2015). For example, procurement auctions often exhibit a "home bias," with buyers favoring local investments about which they have private information (Trionfetti 2000, Brülhart and Trionfetti 2001, Shingal 2015).

Third, shortening interfactory spans lessens travel damage, as one third-party logistics provider explains: "From travel distance to changes in road conditions and elevation, transportation can have wide-ranging effects on the level of product damage. Stiffer trailers combined with rough roads can magnify the compressive stress on boxes by a factor of ten" (Bodenheimer 2014).

4. Data

Our sample spans 1999 to 2014. It includes 28,557 supply chains, 685 car model years, and 47 component types. We create it from four data sets: (1) SupplierBusiness's Who Supplies Whom, (2) Bureau van Dijk's Orbis, (3) NHTSA's SaferCar, and (4) IHS Automotive's PolkInsight.

4.1. Who Supplies Whom

SupplierBusiness's Who Supplies Whom (WSW) data set provides the buyer-seller supply chain links and downstream factory locations (see Figure 1 and Table 1). A division of IHS Automotive, SupplierBusiness claims "all the major automotive manufacturers and over 250 of the world's leading suppliers" as customers.

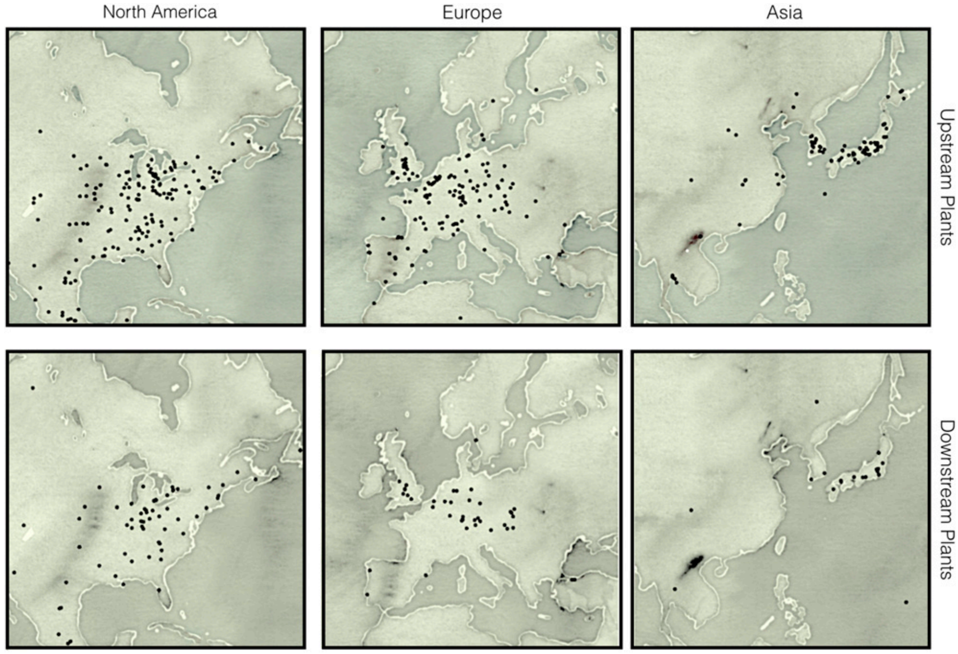
The WSW data set provides supply chain snapshots of specific model-year versions, reporting which supplier made which component. Spanning 1994 to 2015, the data set includes 30 auto groups, 68 automakers, 877 car models, and 1,765 suppliers from over 40 countries.

WSW is similar to Compustat's customer segment data and Bloomberg's supply chain sample. Whereas the Compustat and Bloomberg samples are firm level, the WSW sample is at the product level. For example, Compustat reports that American Axle & Manufacturing (AAM) received $1.56 billion from General Motors in 2009. In contrast, WSW reports that AAM supplied General Motors with differential gears for the 2009 Chevrolet (which was produced at the Ingersoll assembly plant in Canada with a 133,566-unit peak production) and the rear drive module, power train assembly, and rear driveshafts for the 2009 Cadillac SRX (which were produced at the Ramos Arizpe assembly plant in Mexico with a 30,578-unit peak production). Compustat and Bloomberg, respectively, report 60,000 and 15,000 dyads across 70 economic sectors—two-level Standard Industry Classification (SIC) codes—whereas WSW reports 171,107 dyads in a single sector.

4.2. Orbis

Bureau van Dijk's Orbis data set reports the factory locations and financials of suppliers (see Figure 1 and Table 2). It describes 80 million European, 40 million American, and 27 million Asian companies. We select the 14,798 suppliers with NAICS industry code 3361

Figure 1. (Color online) Illustration of Factory Locations Across Three Continents



(motor vehicle manufacturing), 3362 (motor vehicle body and trailer manufacturing), or 3363 (motor vehicle parts manufacturing) that have more than 100 employees. In total, we observe the latitude and longitude of 26,375 supplier factories in 7,639 cities and 135 countries. For example, Orbis reports that supplier Wagon Automotive S.R.L. has factories in Fiano and Chivasso, Italy.

4.3. SaferCar

NHTSA's SaferCar data set provides product-level defect reports (see Table 3). NHTSA's Office of Defects Investigation compiles these defect reports from Vehicle Owner Questionnaires that ask (1) for a description of the component failure; (2) for the car's vehicle identification number, make, model, assembly plant, and year; (3) whether the defect led to a crash, fire, death, or injury; and (4) for the car's speed and the mileage on the odometer. The SaferCar data set includes 976,062 defect reports spanning 21 years and 177,861 auto products.

4.4. PolkInsight

PolkInsight's U.S. National Level Vehicles-in-Operation data set reports vehicle registrations from each state's department of motor vehicles (see Table 4). For example, the sample reports 84,517 registered 2007 Chevy Equinoxes in 2008, 83,975 in 2009, and 83,637 in 2010.

5. Variable Construction

5.1. Dependent Variable

Our primary dependent variable is *Defect Rate*, the total number of defect reports divided by the total number of vehicle-year registrations. For example, 2005 Ford Escape brakes have a rate of 370/6,644,125 = 50.57

Table 1. Sample Observations from the Who Supplies Whom Data Set

Group	Brand	Model year	Module	Component	Assembly plant	Supplier	Peak production	Generation
Honda	Acura	2013 MDX	Fuel tank system	Fuel tank system	United States: Lincoln	Yachiyo of Alabama	74,249	3
BMW Group	BMW	2009 Z4	Steering column/ shaft	Steering shaft sleeve	Germany: Regensburg	Draxlmaier Group	30,651	2
VW Group	Audi	2003 A8	Body side moldings	Body side [outer]	Germany: Neckarsulm	Novelis	22,468	2
Daimler	Mercedes-Benz	2012 A-Class	Air vent/duct	Air duct system	Germany: Rastatt	Rehau AG + CO	175,438	3
Fiat	Alfa Romeo	2008 Mito	Clutch facings	Clutch facings	Italy: Mirafiori	Valeo	57,756	1
Toyota	Toyota	2007 Sequoia	Clutch	One-way clutch	United States: Princeton	Borgwarner	49,683	2

Table 2. Sample Observations from the Orbis Data Set

Supplier	Year	Country	City	Assets	Employees	Patents
Sistemi Sospensioni S.P.A.	2013	Italy	Corbetta	361,087	1,298	34
Sistemi Sospensioni S.P.A.	1999	Poland	Bielsko-Biala	361,087	1,298	34
Sistemi Sospensioni S.P.A.	2012	Italy	Torino	361,087	1,298	34
Qingdao Fangxin Jiacheng Auto Decorations Co., Ltd.	2011	China	Qingdao	24,503	250	0
Limited Liability Company TKM-Service	2003	Russian Federation	Saint Petersburg	1,904	143	0
Wagon Automotive S.R.L.	2010	Italy	Fiano	68,267	207	2
Wagon Automotive S.R.L.	2009	Italy	Chivasso	68,267	207	2

defects per million vehicle-year registrations: SaferCar reports 370 defects for this part and PolkInsight reports 6,644,125 vehicle-year registrations for this model.

5.2. Independent Variable

Our primary independent variable is *Distance*, the geographic separation between the upstream factory and the downstream assembly plant.[1] To construct this variable, we identify the following:

- The location of the vehicle model's assembly plant. WSW lists assembly plant names, which we translate to latitudes and longitudes with the MapQuest API. For example, WSW reports that GM produced the 2007 Buick Enclave in its Lansing Delta plant, and MapQuest positions this plant at coordinates (42.69°N, –84.67°W).
- The location of the supplier's factories. Orbis lists the city of each component factory, which we translate to latitudes and longitudes with the MapQuest API. For example, Orbis reports that Uniwheels Production has a plant in Stalowa Wola, Poland, and MapQuest positions this plant at coordinates (50.57°N, 22.03°E).

We observe which downstream plant assembled which car model and which supplier produced which component, but not which supplier plant produced which component. When the supplier has multiple factories, we define *Distance* as the separation between the assembly plant and the closest supplier factory. In Section 10, we show that dropping the 13% of suppliers with multiple factories yields the same results.

5.3. Control Variables

5.3.1. Product Control Variables.

- *Luxury*: a dummy variable indicating a luxury car model, from the U.S. Consumer Guide classification
- *Generation*: the car model's development stage, from WSW
- *Subcomponents*: the number of subcomponents that make up the auto component, from WSW

Table 3. Sample Observations from the SaferCar Data Set

Manufacturer	Brand	Model year	Date	Component	Problem description
General Motors LLC	Saturn	2006 Vue	April 1, 2014	Steering	"I had to replace the ignition switch a couple . . ."
Mazda Motor Corp	Mazda	1994 Protege	January 14, 2004	Seat belts:front	"Passenger side active shoulder belt seat belt . . ."
General Motors LLC	Chevrolet	2005 Avalanche 1500	March 20, 2011	Electrical system	"Speedometer sticks, don't work and shows wrong. . ."
Starcraft RV, Inc.	Starcraft	2003 Camping Trailer	October 6, 2005	Tires:tread/belt	"I sustained a complete blowout on a left rear. . ."
Forest River, Inc.	Forest River	2006 Cardinal	October 8, 2008	Tires	"Our camper, 2006 Forest River Cardinal, came. . ."
General Motors LLC	Buick	1995 Century	June 23, 2003	Air bags: frontal	"The vehicle suddenly accelerated, and consume. . ."
Ford Motor Company	Ford	2005 Freestyle	June 9, 2014	Structure:body	"Tl: the contact owns a 2005 Ford Freestyle. . ."

Table 4. Sample Observations from the PolkInsight Data Set

Model year	Make	Model	Segment	2008	2009	2010	. . .
2007	Chevrolet	Equinox	Nonluxury CUV	84,517	83,975	83,637	. . .
2008	Chevrolet	Equinox	Nonluxury CUV	79,696	87,541	87,050	. . .
2009	Chevrolet	Equinox	Nonluxury CUV		31,525	31,509	. . .
2010	Chevrolet	Equinox	Nonluxury CUV			135,519	. . .

Table 5. Supply Chain Statistics by Car Brand

	Production		Countries		Plants	
	Start	End	Upstream	Downstream	Upstream	Downstream
Acura	2006	2013	13	2	44	5
Audi	1999	2014	26	3	151	5
BMW	1999	2014	25	3	154	5
Buick	2007	2011	16	2	53	3
Cadillac	2002	2014	21	3	105	5
Chevrolet	2002	2014	21	5	109	17
Chrysler	2003	2010	19	4	95	6
Dodge	2002	2012	18	2	117	10
Fiat	2007	2010	12	2	40	2
Ford	1999	2014	24	6	177	15
GMC	2001	2013	15	3	66	4
Honda	2001	2013	18	5	125	10
Hummer	2005	2005	6	1	7	1
Hyundai	2000	2014	16	5	59	9
Infiniti	2004	2012	9	1	18	2
Jaguar	1999	2009	17	1	71	2
Jeep	2004	2010	15	2	75	4
Kia	2009	2012	11	4	24	4
Land Rover	1999	2012	19	1	61	2
Lexus	2009	2012	10	2	20	3
Lincoln	2002	2012	14	2	71	3
Maruti Suzuki	2002	2007	7	1	17	2
Mazda	1999	2014	14	3	38	4
Mercedes	1999	2014	23	3	125	5
Mercury	2007	2008	12	1	32	2
Mitsubishi	2005	2012	9	2	38	2
Nissan	2003	2014	18	4	119	6
Pontiac	2002	2008	14	2	63	4
Porsche	1999	2012	19	1	83	3
Saab	2005	2005	9	1	27	1
Saturn	2006	2007	14	2	57	4
Subaru	2008	2014	14	2	40	3
Suzuki	1999	2006	10	2	34	2
Toyota	2001	2013	20	9	118	17
Volvo	1999	2012	23	3	82	3
VW	1999	2014	23	6	117	7

- *Volume*: car model sales, from Wards' Auto InfoBank data set

5.3.2. Supplier Control Variables.

- *Employees*: the employee count, from Orbis
- *Assets*: the aggregate assets reported on the supplier's balance sheet, from Orbis
- *Patents*: the number of patents held by the supplier, from Orbis
- *International*: a dummy indicating that the upstream and downstream factories are in different countries

5.3.3. Fixed Effects. We include fixed effects to control for the car model, component type, production year, upstream factory country, and downstream factory country.

Table 5 provides descriptive statistics of the car brands, Table 6 provides summary statistics of the control variables, and Table 7 shows the correlation between variables.

6. Baseline Analysis

6.1. Identification Problem: Selection Endogeneity

We only observe a supplier when a manufacturer chooses to source from it. And the probability that a manufacturer chooses a particular supplier depends on both its distance and quality. Thus, the supplier-selection

Table 6. Summary Statistics: Covariates

Covariate	Mean	SD
Distance	3,821	3,788
Assets	2,898,756	4,320,986
Employees	27,586.8	21,174.6
Generation	3.74	2.43
Luxury	0.34	0.48
International	0.84	0.37
Patents	17	57.8
Volume	62,848	87,986
Subcomponents	43.31	145.83

Table 7. Correlation Coefficients for Independent Variables

	Assets	*Employees*	*Generation*	*Luxury*	*Subcomponents*	*Patents*	*Volume*	*International*
Assets	1	—	—	—	—	—	—	—
Employees	0.681	1	—	—	—	—	—	—
Generation	0.014	0.015	1	—	—	—	—	—
Luxury	–0.040	–0.031	–0.154	1	—	—	—	—
Subcomponents	–0.035	–0.041	0.050	0.026	1	—	—	—
Patents	0.232	0.245	0.002	0.058	0.035	1	—	—
Volume	0.023	0.026	0.290	–0.405	0.100	–0.051	1	—
International	–0.279	–0.217	–0.005	0.065	0.078	0.065	–0.084	1

process can correlate with the distance and quality of our sample's suppliers. To control for this potential selection bias, we jointly model the supplier-selection probabilities and the component defect rates, in the fashion of Dahl (2002).

6.2. Model

6.2.1. Supplier Selection. Let p be the part—for example, the suspension system of the 2011 Jeep Grand Cherokee—and $\mathcal{S}_p$ the set of suppliers capable of producing this part—for example, the 312 suspension system suppliers in our sample. The value of sourcing part p from supplier $s \in \mathcal{S}_p$ is

$$\mathit{Value}_{ps} = \underbrace{\alpha^F_{ps}}_{\text{fixed effects}} + \underbrace{\alpha^C_p \mathit{Controls}_{ps}}_{\text{control variables}} + \underbrace{\sum_{i=1}^{m+n} \alpha^D_{pi} \mathfrak{B}\left(\mathit{Distance}_s, i\right)}_{\text{distance splines}} + \underbrace{\epsilon_{ps}}_{\substack{\text{unobserved}\\ \text{value shock}}}, \tag{1}$$

where ϵ_{ps} has a type I extreme value distribution with PDF dF and $\mathfrak{B}\,(\mathit{Distance}_s, i)$ is the ith basis function of a degree n B-spline with m interior knots spaced according to *Distance*'s quantiles. The parameters $(\alpha^D_{p1}, \ldots, \alpha^D_{p(m+n)})$ are "Dahl coefficients."

6.2.2. Defect Rates. Sourcing part p from supplier s yields

$$\mathit{Defect\ Rate}_{ps} = \underbrace{\beta^F_{ps}}_{\text{fixed effects}} + \underbrace{\beta^C \mathit{Controls}_{ps}}_{\text{control variables}} + \underbrace{\beta^D \mathit{Distance}_s}_{\text{distance effect}} + \underbrace{\eta_{ps}}_{\substack{\text{unobserved}\\ \text{quality shock}}}. \tag{2}$$

6.2.3. Selection Correction. The automaker chooses the highest-valuation supplier. So we only observe $s^* = \arg\max_{s\in\mathcal{S}_p} \mathit{Value}_{ps}$. This censoring can lead to correlation between covariate $\mathit{Distance}_{s^*}$ and error η_{ps^*} in our sample because the selection process influences both: $\mathit{Distance}_{s^*}$ influences Value_{ps^*} directly, and η_{ps^*} influences Value_{ps^*} indirectly via its correlation with ϵ_{ps^*}. To account for this endogeneity, we factor in the expected quality shock conditional on winning the contract. We suppose that η_{ps^*} has conditional distribution $dF(\eta_{ps^*} \,|\, \epsilon_{ps^*})$. The probability of the automaker choosing supplier $s \in \mathcal{S}_p$ is

$$\Psi_{ps} = \frac{\exp\left(\alpha^F_{ps} + \alpha^C_p \mathit{Controls}_{ps} + \sum_{i=1}^{m+n} \alpha^D_{pi}\mathfrak{B}\left(\mathit{Distance}_s, i\right)\right)}{\sum_{s'\in\mathcal{S}_p}\left(\exp\left(\alpha^F_{ps'} + \alpha^C_p \mathit{Controls}_{ps'} + \sum_{i=1}^{m+n}\alpha^D_{pi}\mathfrak{B}\left(\mathit{Distance}_{s'}, i\right)\right)\right)}. \tag{3}$$

Variable ϵ_{ps^*} has conditional PDF

$$dF(\epsilon_{ps^*} \,|\, s^* = \arg\max_{s\in\mathcal{S}_p} \mathit{Value}_{ps}) = dF\left(\epsilon_{ps^*} + \log(\Psi_{ps^*})\right). \tag{4}$$

Accordingly, given that supplier s^* won the procurement contract, variable η_{ps^*} has conditional expectation

$$\mathbb{E}\left(\eta_{ps^*} \,|\, s^* = \arg\max_{s\in\mathcal{S}_p} \mathit{Value}_{ps}\right) = \iint \eta_{ps^*} \cdot dF\left(\eta_{ps^*} | \epsilon_{ps^*}\right) \cdot dF\left(\epsilon_{ps^*} + \log(\Psi_{ps^*})\right). \tag{5}$$

Term $\mathbb{E}(\eta_{ps^*} \,|\, s^* = \arg\max_{s\in\mathcal{S}_p} \mathit{Value}_{ps})$ depends on the first- and second-stage covariates through Ψ_{ps^*}: the selection probability is a sufficient statistic for the selection bias. Thus, we can approximate $\mathbb{E}(\eta_{ps^*} \,|\, s^* = \arg\max_{s\in\mathcal{S}_p} \mathit{Value}_{ps})$ arbitrarily closely with a B-spline function of Ψ_{ps^*}. Following Dahl (2002), we add these B-splines directly to our defect-rate equation:

$$\mathit{Defect\ Rate}_{ps^*} = \beta^F_{ps^*} + \beta^C \mathit{Controls}_{ps^*} + \beta^D \mathit{Distance}_{ps^*} + \hat{\mathbb{E}}[\eta_{ps^*} \,|\, s^* = \arg\max_{s\in\mathcal{S}_p} \mathit{Value}_{ps}] + \hat{\eta}_{ps^*} \tag{6}$$

where

$$\mathbb{E}[\eta_{ps^*} \,|\, s^* = \arg\max_{s\in\mathcal{S}_p} \mathit{Value}_{ps}] = \sum_{i=0}^{m+n} \gamma_i \mathfrak{B}^{\Psi}\left(\Psi_{ps^*}, i\right)$$

and

$$\hat{\eta}_{ps^*} = \eta_{ps^*} - \hat{\mathbb{E}}[\eta_{ps^*} \,|\, s^* = \arg\max_{s\in\mathcal{S}_p} \mathit{Value}_{ps}].$$

The term $\mathfrak{B}^{\Psi}(\Psi_{ps^*}, i)$ is the ith basis function of a degree n B-spline with m interior knots spaced according to the quantiles of Ψ_{ps^*}. The expectation $\mathbb{E}[\eta_{ps^*} \mid s^* = \arg\max_{s\in\mathcal{S}_p} Value_{ps}]$ is conditional on

$$\mathbb{E}[Defect\ Rate_{ps^*}] = \beta^F_{ps^*} + \beta^C Controls_{ps^*} + \beta^D Distance_{ps^*}.$$

Thus, our new error term, $\hat{\eta}_{ps^*}$, is orthogonal to $\mathbb{E}[Defect\ Rate_{ps^*}]$'s constituent parts. As a result, we wash out the defect rate variation that correlates with unobserved quality attributes.

We estimate our model in two steps:

1. We estimate the supplier valuation coefficients with a set of multinomial logistic regressions (one per component type) that maximize the likelihood of the observed supplier configuration. We set the choice probability estimates, $\hat{\Psi}_{ps}$, to the regressions' fitted values. We set $n = m = 3$, characterizing the relationship between the distance to a supplier and the value of that supplier with four third-degree-polynomial splines.
2. We estimate expression (6) using $\hat{\Psi}_{ps}$. Again, we set $m = n = 3$ in the spline sum.

We use a multiplicative specification, taking the logarithm of the continuous variables, because they are positive and skewed. And we calculate standard errors with the bootstrap, sampling supply chain links randomly (Newey 1984). In Section 11, we show that our results hold under alternative specifications.

6.3. Estimates

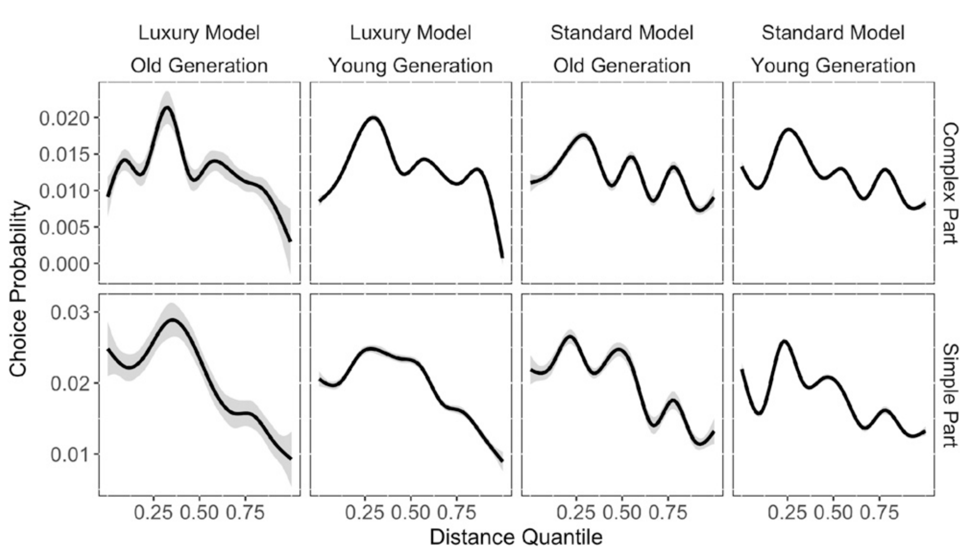
6.3.1. Supplier Selection. Figure 2 shows how supplier-choice probability estimates vary with supply chain distance.[2] Overall, the chosen suppliers have an average selection probability of 0.13, and the unchosen suppliers have an average selection probability of 0.014. Figure 3 depicts the empirical fit of our supplier selection model. When pooled, our multinomial logistic regressions have a McFadden R-squared of 0.30, both in and out of sample (which suggests we did not overfit the model).

6.3.2. Defect Rates. Table 8 shows that the relationship between *Distance* and *Defect Rate* is positive and statistically significant across all specifications. Our 0.0083 elasticity estimate implies that scaling supply chain distance by one order of magnitude increases expected defect rates by $100 \cdot (100^{0.0083} - 1) = 3.90\%$.

To interpret these results, we conduct two counterfactual analyses: a ceteris paribus counterfactual and a mutatis mutandis counterfactual (see Figure 4). Both use our statistical model to anticipate what the average *Defect Rate* would have been had the *Distance* variable been that of the nth-next-closest qualified part supplier. The ceteris paribus counterfactual changes only the *Distance* variable, whereas the mutatis mutandis counterfactual changes all the supplier variables. Table 4 depicts these counterfactuals. The ceteris paribus and mutatis mutandis analyses suggest that replacing the current set of suppliers with the next-closest set would increase expected defect rates by 1.57% and 2.17%, respectively.

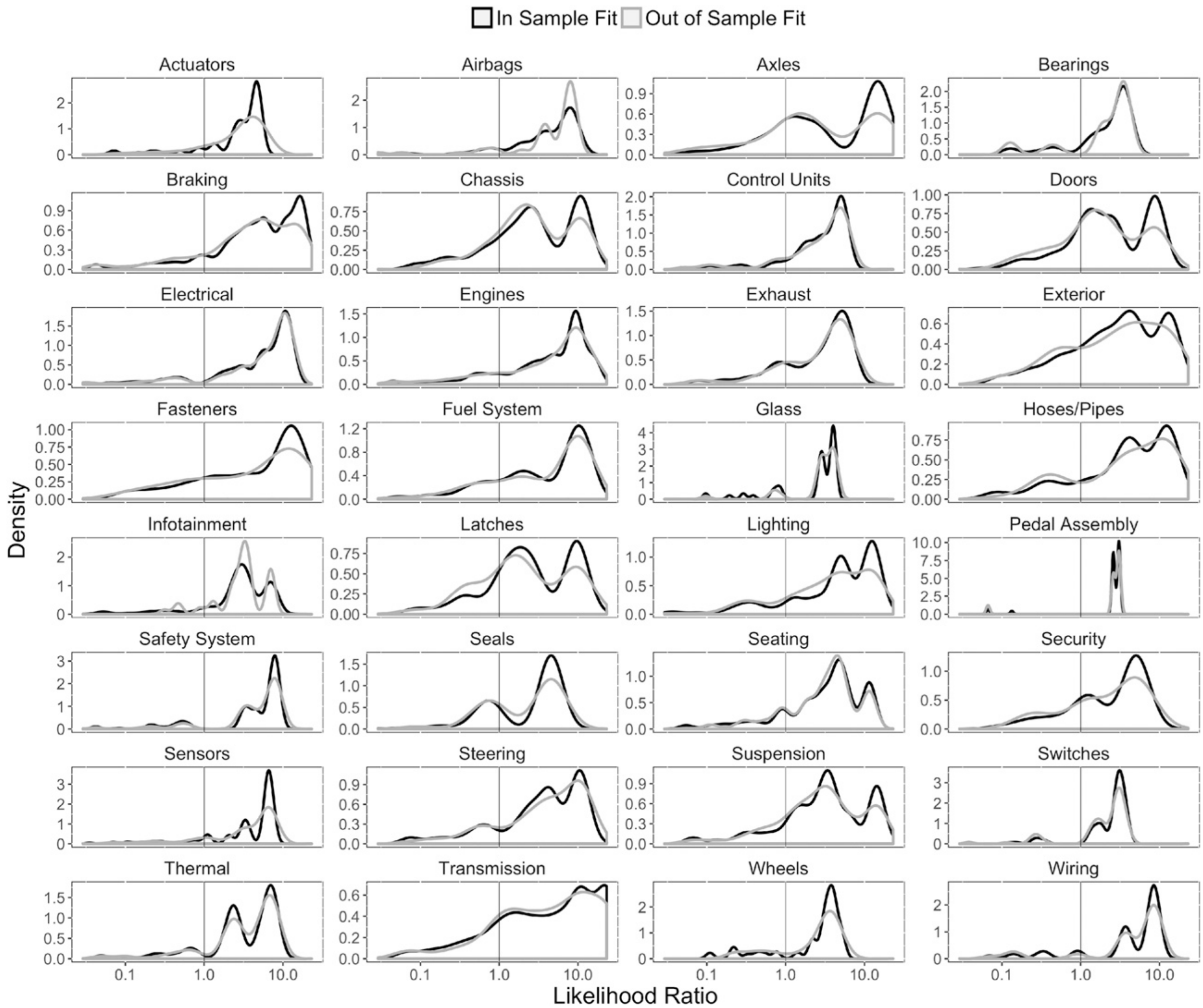
Figures 5 and 6 plot the relationship between supplier proximity and product quality by component type and brand. The correlation is significantly positive across 13 component types and 10 brands

Figure 2. Selection Model Fit



Notes. These plots depict the fitted values of LOESS regressions of choice probability estimates of the selection probability, $\hat{\Psi}_{ps}$, on the quantiles of $Distance_{ps}$. We run eight regressions, dividing the sample by (1) whether the car model is luxury or standard, (2) whether the car model generation number is larger or smaller than the median, and (3) whether the component count is larger or smaller than the median. We depict the regression curves' 95% confidence intervals with gray bands.

Figure 3. Out-of-Sample Test



Notes. These plots illustrate the empirical fit of our supplier selection model. In the spirit of McFadden's R^2 calculation, we graph the likelihood ratio of the selected supplier's choice probability under our model and under the null model in which all qualified suppliers have an equal chance of being selected. This likelihood ratio exceeds one when our model outperforms the null model, that is, when *Likelihood*(*Fit*) > *Likelihood*(*NullModel*). The black curves depict the in-sample fit-in which we use all the data to train and fit the model—and the gray curves depict the out-of-sample fit—in which we train the model on half of our data and fit it on the other half.

and is significantly negative across one component type and two brands.

This subsample analysis reveals that different component types and car brands respond differently to supply chain distance. To explain this heterogeneity, we study the factors that moderate the effect of distance. Table 9 indicates that distance is more detrimental to product quality when

1. *The component resides in a high-end (luxury) product.* The elasticity of defect rates to supply chain distance is larger in luxury model parts than in standard model parts. High-end products require quality materials and customized components that must last longer and meet higher performance standards.

2. *The component has many subcomponents.* The elasticity of defect rates to supply chain distance is larger when the component comprises multiple subcomponents. Complexity necessitates proximity: producing a complex component requires more oversight and monitoring than producing a stand-alone part.

3. *The product is new.* The elasticity of defect rates to supply chain distance is significantly larger in early-generation model parts than in late-generation model parts. Learning by doing counteracts the problems that distance creates.

4. *The supplier invests little in research and development (R&D).* The elasticity of defect rates to supply chain distance is larger when the supplier has few production patents. Less oversight of vendors is needed when they do research and development.

5. *Production is outsourced to another country.* An international mile counts more than a domestic mile. The

Table 8. Defect-Rate Model Estimates

	1	2	3	4	5	6	7	8	9
Distance	0.83***	0.85***	0.84***	0.85***	0.83***	0.82***	0.81***	0.80***	0.80***
	(0.13)	(0.12)	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)
Luxury	—	7.09***	7.09***	6.47***	3.63***	2.83***	2.84***	2.84***	2.83***
	—	(0.53)	(0.53)	(0.60)	(0.59)	(0.58)	(0.58)	(0.58)	(0.58)
International	—	—	0.04	0.00	–0.18	0.05	0.05	0.04	0.09
	—	—	(0.62)	(0.62)	(0.61)	(0.61)	(0.61)	(0.61)	(0.61)
Generation	—	—	—	0.65**	–0.35	–1.24***	–1.24***	–1.23***	–1.24***
	—	—	—	(0.30)	(0.30)	(0.31)	(0.31)	(0.31)	(0.31)
Subcomponents	—	—	—	—	5.03***	3.77***	3.77***	3.77***	3.78***
	—	—	—	—	(0.17)	(0.19)	(0.19)	(0.19)	(0.19)
Assets	—	—	—	—	—	20.54***	20.54***	20.55***	20.53***
	—	—	—	—	—	(1.43)	(1.43)	(1.43)	(1.43)
Employees	—	—	—	—	—	—	–0.03	0.00	0.08
	—	—	—	—	—	—	(0.11)	(0.12)	(0.12)
Patents	—	—	—	—	—	—	—	–0.09	–0.05
	—	—	—	—	—	—	—	(0.10)	(0.10)
Volume	—	—	—	—	—	—	—	—	–0.20**
	—	—	—	—	—	—	—	—	(0.09)
Dahl coefficient 1	0.22	0.28	0.28	0.39	–0.15	–0.15	–0.14	–0.11	–0.18
	(1.89)	(1.89)	(1.89)	(1.87)	(1.80)	(1.77)	(1.78)	(1.78)	(1.78)
Dahl coefficient 2	–0.78	–0.76	–0.76	–0.75	–1.01	–0.89	–0.90	–0.85	–0.82
	(1.15)	(1.15)	(1.15)	(1.14)	(1.16)	(1.18)	(1.18)	(1.20)	(1.20)
Dahl coefficient 3	–0.92	–0.73	–0.73	–0.63	–0.90	–0.78	–0.78	–0.73	–0.56
	(1.47)	(1.48)	(1.47)	(1.46)	(1.43)	(1.38)	(1.38)	(1.40)	(1.39)
Dahl coefficient 4	1.82	1.68	1.68	1.68	0.70	1.01	0.99	1.01	1.34
	(2.22)	(2.23)	(2.22)	(2.19)	(2.18)	(2.19)	(2.21)	(2.23)	(2.23)
Dahl coefficient 5	–4.56	–3.83	–3.84	–3.52	–3.28	–3.55	–3.54	–3.48	–3.82
	(3.87)	(3.81)	(3.77)	(3.77)	(3.65)	(3.48)	(3.49)	(3.53)	(3.52)
Dahl coefficient 6	–1.94	–2.36	–2.36	–2.37	–4.56	–3.95	–3.94	–3.69	–2.96
	(5.96)	(5.99)	(6.02)	(6.04)	(5.90)	(5.89)	(5.91)	(5.93)	(5.88)
R^2	0.23	0.23	0.23	0.23	0.24	0.25	0.25	0.25	0.25
VIF	—	1.00	1.35	1.36	1.37	1.37	1.42	1.70	1.92
N	23.79	23.79	23.79	23.79	23.79	23.79	23.79	23.79	23.79

Notes. This table presents estimates of the defect rate coefficients. Each column represents a different regression; we build up to our full specification gradually, adding one control variable at a time. Below, the rows report the regressions' R^2, variance inflation factor (VIF), and sample size N (measured in thousands of observations). For brevity, we do not tabulate the production year, brand, component, or country fixed effects. For illustrative purposes, we scale the estimates by a factor of 100 (e.g., the top-left estimate is actually 0.0100). We present bootstrapped standard errors. The sample horizon is 1999–2014.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

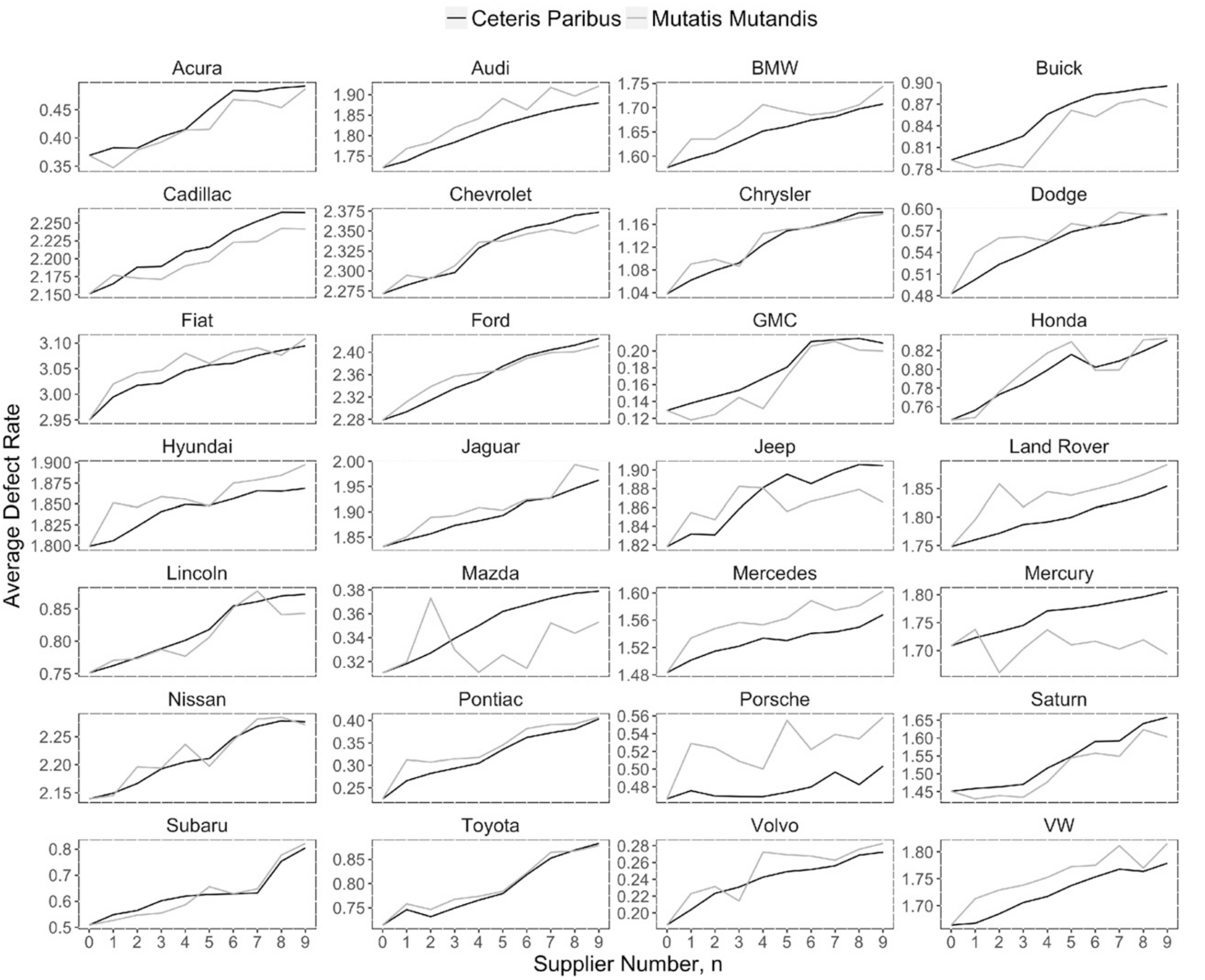
elasticity of defect rates to supply chain distance is larger when the part is produced by a foreign supplier. Crossing borders makes travel arduous and creates cultural, economic, and political barriers to the supply chain.

6. *The model has a large production volume.* Scaling up an operation exposes the fragility of long supply chains. The elasticity of defect rates to supply chain distance increases with production volume. Proximity makes the supply chain more resilient to the stresses of scaling.

7. Factory Relocations

Automakers sometimes move production from one assembly plant to another. We now treat these relocations as exogenous distance shocks. The relocations simultaneously shift the supply chain distance for every auto component in the car by lengthening some and shortening others. For example, in 2010 Toyota relocated the assembly of its tenth-generation Corolla from its NUMMI plant in Fremont, California to Blue Springs, Mississippi, after Governor Haley Barbour "promised Toyota a 20-year corporate tax holiday" (Dawson 2010). The move did not change the Corolla's supplier base (see Table 10) or its production process, because Toyota shipped its "production equipment . . . from the recently shuttered NUMMI plant (now owned by Tesla) in California [to Blue Springs]" (Dawson 2010). But the shift did change the Corolla's supply chain distances: for example, the distance to Gentex's interior mirror factory in Zeeland, Michigan, decreased

Figure 4. Counterfactual Analysis



Notes. This figure depicts the results of our two counterfactual analyses. The ceteris paribus counterfactual estimates what the average defect rate would have been had the supply chain distance variable been that of the *n*th-next-closest qualified part supplier (without any other variables changing). The mutatis mutandis counterfactual, in contrast, allows other variables to vary contiguously with *Distance*. Then = 0 values are the current defect rates.

from 2,188 miles to 721.5 miles, whereas the distance to SIKA Automotive's factory in Santa Fe Springs, California, increased from 342.4 miles to 1,812.3 miles (see Figure 7).

Our data set reports 79 such plant shifts. For instance,

• Mercedes-Benz's production moved to Tuscaloosa after the government of Alabama promised "more than $200 million in job training, tax breaks, and other incentives and even offered to rename a section of Interstate 20/59 the Mercedes-Benz Autobahn" (Woodruff and Templeman 1993).

• Honda moved production of its 2011 Accord MDX sport-utility model from Lincoln, Alabama, to East Liberty, Ohio, because "it need[ed] to free up space both for the redesigned 2016 Honda Pilot SUV, which is likely to see sales climb for its new iteration, and for the Honda Ridgeline truck, production of which will resume next year with a redesigned version" (Eaton 2015).

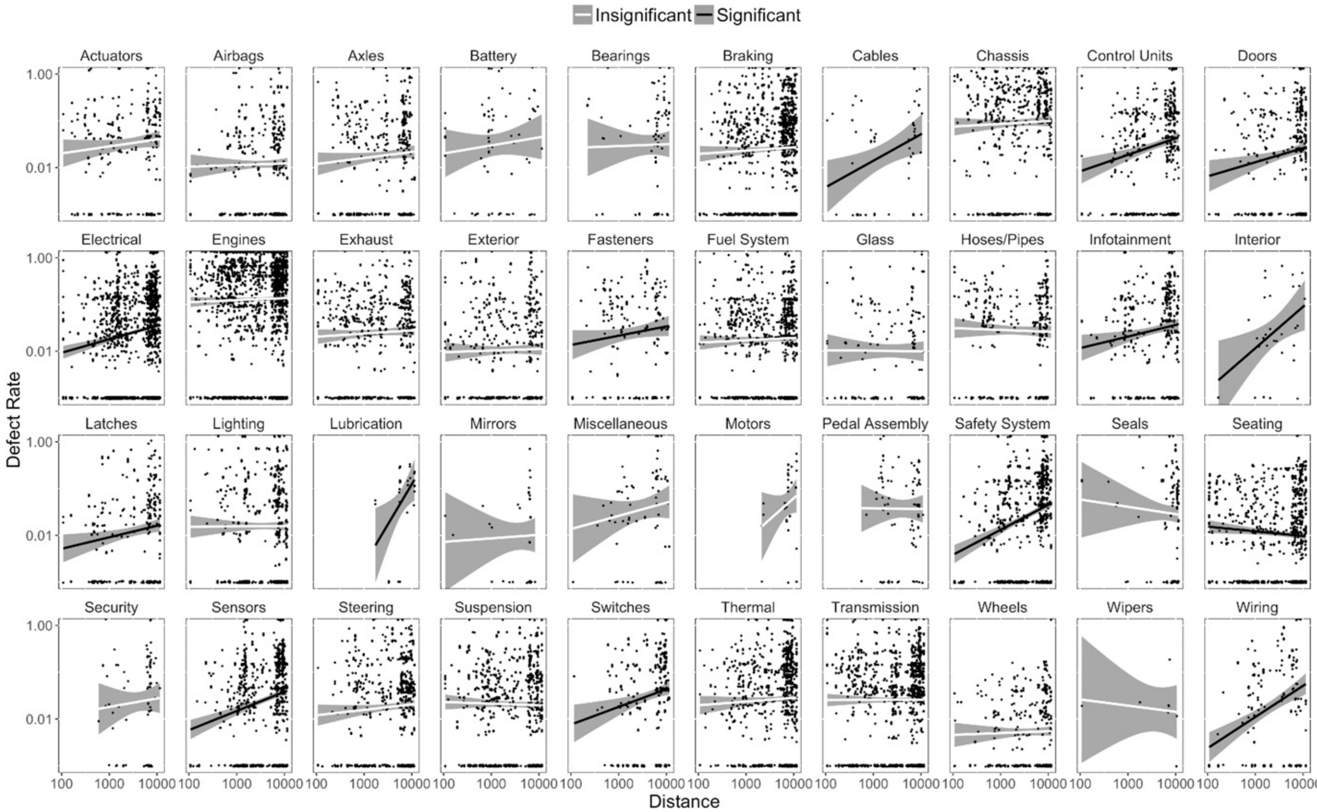
• Ford moved the production of its 2011 Explorer from Louisville, Kentucky to Chicago, Illinois, after investing "$400 million in its Chicago assembly and metal-stamping plants" (Pete 2015).

Collectively, the 79 relocations affected 1,203 sourcing relationships, spanning 2,451 products. The plant moves repositioned vehicle assembly closer to 553 supplier factories and farther from 650 supplier factories. The supply chains that shortened did so by 3,105.6 km, on average, and those that lengthened did so by 3,184.2 km. The relocations divide our sample into three groups:

1. A treated group of suppliers that experienced a distance increase following relocation.

2. A treated group of suppliers that experienced a distance decrease following relocation.

Figure 5. Supply Chain Distance and Product Quality by Component Type



3. A control group of suppliers that didn't experience a relocation.

The two treated groups serve as benchmarks for one another; for instance, if the factory workers in the new facility are more skilled, the relocation should impose a positive shock on quality two treated groups. The only factor that changes disparately with the relocation is supply chain distance. And Table 11 reports that the relocations improve quality when distances shorten and harm quality when distances increase.

We regress the change in the defect rates on the change in supply chain distances with the following equation:

$$\Delta DefectRate_{ps} = \beta_{FE,ps} + \beta_x' x_{ps} + \Delta Distance_{ps} + \epsilon_{ps}.$$

The dependent variable is the defect-rate change across two production years of a model component, and the primary independent variable is the corresponding distance change. For example, we find (1) a defect rate of 0.23 for the vibration controls of the 2009 Corolla (manufactured by Sika and assembled by Toyota in Fremont, California) and (2) a defect rate of 0.27 for the vibration controls of the 2010 Corolla (manufactured by Sika and assembled by Toyota in Blue Springs, MS); accordingly, the dependent variable is 0.04 ($\Delta DefectRate$ = 0.27 − 0.23). We use the same covariates as before, except we have replaced the *International* dummy with two new dummies: (1) *Out-to-In*, which equals 1 when the automaker relocates assembly from abroad to the supplier's home country, and (2) *In-to-Out*, which equals 1 when the automaker relocates assembly from the supplier's home country to abroad.

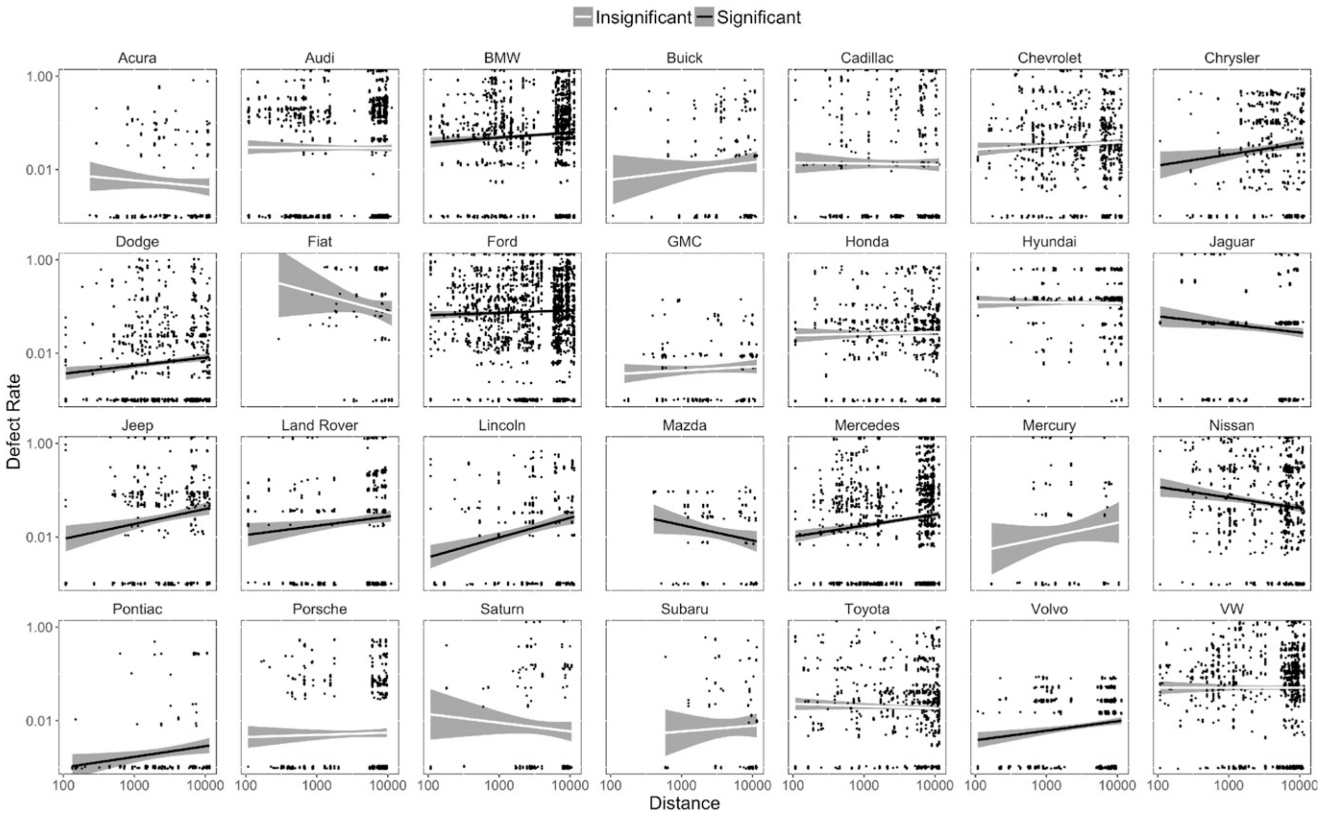
Table 12 reports the results. We find a significant positive relationship between supply chain distance and defect rates. Further, we find defect rates decrease when the automaker relocates to the supplier's country.

8. Proximity and Quality Improvement

Our data reveal that relationships are persistent in the automotive industry and that suppliers and automakers work together to improve subsequent versions of a car model. Consider the Audi A6's alternator, supplied by Valeo, whose defect rate fell by 13% between its 2001 and 2005 models. A web search reveals that Valeo upgraded the alternator's specifications on two occasions, changing its amperage, weight, and dimensions.

Could geographic distance propel this rate of quality improvement? In Section 3, we argued it does, because proximity facilitates monitoring, collaboration, and face-to-face interactions—especially in the automotive industry, where 60%–80% of auto parts are redesigned with every iteration of a car model (Clark and Fujimoto 1991, p. 149) and where tier-1 components are often tailored to specific models (Dyer 2000). Thus, far-flung suppliers should improve quality more slowly.

Figure 6. Supply Chain Distance and Product Quality by Car Brand



Notes. The two plots above depict *Defect Rate*, our primary dependent variable, and *Distance*, our primary independent variable. We use a log-log scale, winsorizing the tails at 3%. We eliminate component types with fewer than 40 observations. We depict the fitted regression curve with a line; this line is black when the slope is significantly different from zero, and white otherwise. We depict the regression curves' 95% confidence intervals with gray bands.

To test this hypothesis, we estimate how the evolution of defect rates depends on supply chain distance with a sample of 1,783 product-model-supplier triplets observed across subsequent model-year iterations. We regress *Rate Change* on *Distance*, where *Rate Change* is the change in the defect rate between the two model-year iterations. Table 13 reports that the intercept is negative, which indicates that quality systematically improves over time. But the *Distance* coefficient is positive, which indicates that this improvement is slower when the supplier and automaker are farther apart.

This result is supported by anecdotal evidence. From managerial interviews, we learned that automakers rarely get it right from the get-go. It is standard practice in automotive product development to maintain a "queue" of known issues and complaints to eventually get around to. After several model-year iterations, some products improve; others do not. Spindelndreier et al. (2015) believe that this disparity could be, in part, driven by supply chain proximity:

> [A component supplier] had two plants that were very close to a Toyota facility . . . . And those plants were strikingly better than the average in their processes . . . because Toyota's spending a lot of effort training and teaching those guys. . . . I mean, it's just, if you're down the road with your plant, if you're having problems you deal with each other and you try to solve it. And you start working, collaborating very closely. You start maybe discussing even a small deviation before it becomes an actual product. If you're much farther away you likely will not do that.

9. Travel Time

The value of proximity may depend not only on how far apart the two factories are but also on how quickly the supply chain partners can travel from one factory to the other. We reestimate our results with an alternative distance metric: the average travel time between upstream and downstream factories. To estimate this effect, we focus on those pairs of factories that can be connected through ground routes (see Figure 8).

We measure the average interfactory travel time with *Travel Time* = *Travel Distance*/*Average Travel Speed*. *Travel Distance* is the distance of the most efficient driving route between the factories and *Average Travel Speed* factors in slowdowns caused by stoplights, bridges, tolls, etc. We derive both metrics from OpenStreetMap.com, procuring all ground travel routes and the average weekly travel speed across every directed kilometer segment

Table 9. Defect-Rate Model Estimates with Interactions

	1	2	3	4	5	6	7	8	9
Distance	0.37**	0.73**	0.17***	0.31***	0.20***	0.38***	0.11**	–0.10**	0.44***
	(0.15)	(0.31)	(0.34)	(0.34)	(0.51)	(1.74)	(1.80)	(1.78)	(1.71)
Distance · Luxury	1.02***	0.95***	0.92***	0.93***	0.94***	1.09***	1.06***	1.06***	1.04***
	(0.21)	(0.22)	(0.22)	(0.23)	(0.23)	(0.22)	(0.22)	(0.22)	(0.22)
Distance · International	—	1.47***	1.46***	1.45***	1.38***	1.42***	1.40***	1.45***	1.41***
	—	(0.37)	(0.37)	(0.37)	(0.38)	(0.38)	(0.39)	(0.39)	(0.39)
Distance · Generation	—	—	—	0.03	–0.04**	–0.14***	–0.15***	–0.14***	–0.16***
	—	—	—	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.2)
Distance · Subcomponents	—	—	—	—	0.34**	0.28*	0.27*	0.27*	0.29*
	—	—	—	—	(0.13)	(0.15)	(0.15)	(0.15)	(0.15)
Distance · Assets	—	—	—	—	—	1.46	1.47	1.46	1.4
	—	—	—	—	—	(1.69)	(1.68)	(1.68)	(1.69)
Distance · Employees	—	—	—	—	—	—	–0.08	–0.17	–0.13*
	—	—	—	—	—	—	(0.05)	(0.10)	(0.10)
Distance · Patents	—	—	—	—	—	—	—	0.15*	0.21**
	—	—	—	—	—	—	—	(0.08)	(0.09)
Distance · Volume	—	—	—	—	—	—	—	—	–0.12*
	—	—	—	—	—	—	—	—	(0.07)
Luxury	–5.49***	–4.93***	–4.76**	–4.78**	–4.90**	–6.16***	–5.95***	–5.93***	–5.77***
	(1.81)	(1.89)	(1.88)	(1.93)	(1.93)	(1.94)	(1.89)	(1.89)	(1.88)
International	0.17	–10.11***	–10.10***	–10.08***	–9.47***	–9.76***	–9.55***	–9.87***	–9.35***
	(0.61)	(2.52)	(2.53)	(2.56)	(2.60)	(2.64)	(2.68)	(2.71)	(2.76)
Generation	–1.22***	–1.17***	–1.13***	–1.39	–0.85	–0.01	0.13	–0.03	0.15
	(0.31)	(0.31)	(0.31)	(1.63)	(1.58)	(1.64)	(1.65)	(1.62)	(1.64)
Subcomponents	3.77***	3.76***	3.80***	3.80***	1.07	1.54	1.60	1.64	1.43
	(0.19)	(0.19)	(0.19)	(0.19)	(1.12)	(1.21)	(1.20)	(1.20)	(1.26)
Assets	20.50***	20.52***	20.51***	20.51***	20.52***	8.82	8.75	8.70	9.02
	(1.43)	(1.42)	(1.42)	(1.42)	(1.41)	(6.56)	(6.54)	(6.52)	(6.59)
Employees	0.08	0.12	0.11	0.11	0.11	0.11	0.74	1.54**	1.23**
	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.46)	(0.60)	(0.60)
Patents	–0.05	–0.03	–0.03	–0.03	–0.03	–0.04	–0.06	–1.28*	–1.74**
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.66)	(0.69)
Revenue	–0.20**	–0.21**	–0.21**	–0.21**	–0.21**	–0.21**	–0.20**	–0.19**	0.78
	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.55)
Dahl coefficient 1	–0.20	–0.28	–0.27	–0.27	–0.22	–0.19	–0.20	–0.23	–0.29
	(1.77)	(1.76)	(1.77)	(1.77)	(1.77)	(1.76)	(1.77)	(1.76)	(1.78)
Dahl coefficient 2	–0.82	–0.87	–0.91	–0.91	–1.07	–1.00	–1.04	–1.07	–1.05
	(1.20)	(1.19)	(1.20)	(1.20)	(1.19)	(1.20)	(1.19)	(1.19)	(1.19)
Dahl coefficient 3	–0.57	–0.48	–0.44	–0.44	–0.32	–0.32	–0.37	–0.36	–0.38
	(1.39)	(1.39)	(1.39)	(1.39)	(1.39)	(1.39)	(1.39)	(1.39)	(1.39)
Dahl coefficient 4	1.23	1.00	0.87	0.87	0.50	0.65	0.61	0.63	0.92
	(2.24)	(2.24)	(2.25)	(2.25)	(2.23)	(2.25)	(2.26)	(2.26)	(2.23)
Dahl coefficient 5	–3.44	–3.01	–2.72	–2.72	–2.93	–3.01	–3.12	–3.23	–3.80
	(3.53)	(3.45)	(3.46)	(3.47)	(3.46)	(3.48)	(3.44)	(3.48)	(3.39)
Dahl coefficient 6	–3.57	–4.57	–4.74	–4.75	–4.43	–4.26	–4.38	–4.34	–4.05
	(5.95)	(6.07)	(6.12)	(6.11)	(5.94)	(5.99)	(5.98)	(5.98)	(5.88)
R^2	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
VIF	46.65	62.60	62.60	62.86	63.30	337.43	337.54	337.55	337.75
N	23.79	23.79	23.79	23.79	23.79	23.79	23.79	23.79	23.79

This table presents estimates of our defect rate model, extended to include up to eight additional interaction terms. We present the estimates in the fashion of Table 8. The sample horizon is 1999–2014.

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

in each road. After obtaining the various travel routes, an applied programming interface (API) gives us the shortest traveling route between the two factories. For example, to go from Chicago, Illinois to Columbus, Ohio, we find three routes: (1) via I-65 S and I-70 E (5 hr 8 min; 356 miles), (2) via US-30 E (5 hr 47 min; 327 miles),

Table 10. Random Sample of Toyota's Supplier Base for the 2010 and 2011 Corolla Models

Model	Year	Plant	Generation	Product	Supplier
Corolla	2010	Blue Springs	10	Constant Velocity Joints	NTN
Corolla	2011	Fremont	10	Constant Velocity Joints	NTN
Corolla	2010	Blue Springs	10	Engine Shafts	Kautex Textron
Corolla	2011	Fremont	10	Engine Shafts	Kautex Textron
Corolla	2010	Blue Springs	10	Fasteners/Fixings	Midway Products
Corolla	2011	Fremont	10	Fasteners/Fixings	Midway Products
Corolla	2010	Blue Springs	10	Friction Plates	Borgwarner
Corolla	2011	Fremont	10	Friction Plates	Borgwarner
Corolla	2010	Blue Springs	10	Interior Mirror	Gentex
Corolla	2011	Fremont	10	Interior Mirror	Gentex
Corolla	2010	Blue Springs	10	Lock Systems	TRW
Corolla	2011	Fremont	10	Lock Systems	TRW
Corolla	2010	Blue Springs	10	Vibration Controls	Sika
Corolla	2011	Fremont	10	Vibration Controls	Sika
Corolla	2010	Blue Springs	10	Seals [body]	Sika
Corolla	2011	Fremont	10	Seals [body]	Sika
Corolla	2010	Blue Springs	10	Seat Belts	Takata
Corolla	2011	Fremont	10	Seat Belts	Takata

and (3) via I-90 E (5 hr 55 min; 363 miles). We choose route (1), which, despite being longer than route (2), is the quickest.

The correlation between *Travel Time* and *Distance* is 0.961. Unsurprisingly, Table 14 shows that all our baseline results are identical.[3]

10. Robustness Analysis

Estimates for all robustness tests are shown in Table 15.

10.1. Suppliers with Multiple Factories

We observe the geographic locations of suppliers' factories, but we don't observe which factory produces what product. As a robustness check, we reestimate our results by restricting our sample to suppliers with only one factory (which results in dropping 13% of our observations). Eliminating multifactory suppliers obviates the ambiguity in upstream production location: we know where a supplier does its manufacturing when

Figure 7. (Color online) Geographic Distances for Sika and Gentex, Pre- and Postrelocation

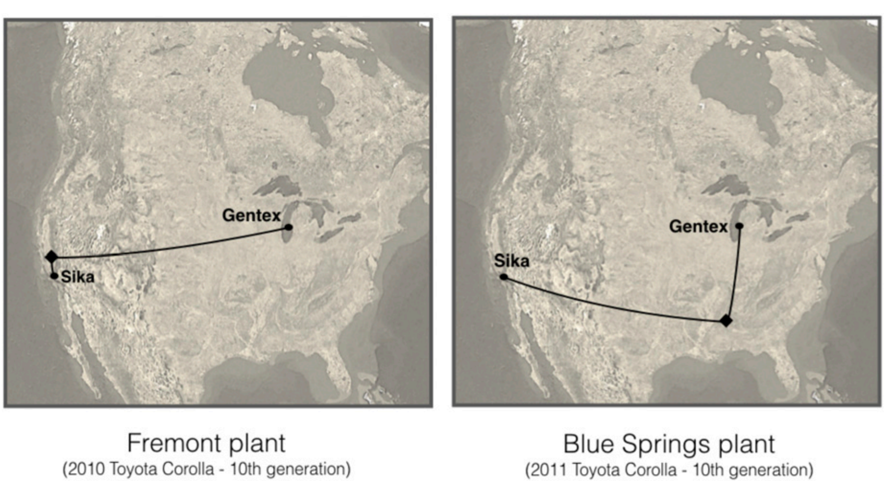


Table 11. Summary Statistics Displaying Changes in *Distance* and *Defect Rate* Postrelocation

	Distance increase	Distance decrease	No change (control)
Δ*Distance (km)*	3,184.23	–3,105.58	0
Δ*Distance (%)*	46.10%	–47.61%	0%
Δ*Defect rate*	0.83	–0.23	–0.07
Δ*Defect rate (%)*	2.10%	–1.11%	–0.16%

Table 12. Secondary Regression Results Using Model Relocations

	1	2	3	4	5	6	7	8	9
ΔDistance	0.85**	0.85**	0.97***	0.98***	0.97***	0.97***	1.00***	1.00***	1.00***
	(0.35)	(0.35)	(0.37)	(0.37)	(0.37)	(0.37)	(0.38)	(0.38)	(0.38)
Luxury	—	0.17***	0.16***	0.10***	0.12***	0.20***	0.19***	0.19***	0.19***
	—	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
In-to-Out	—	—	–0.43	–0.48	–0.47	–0.42	–0.42	–0.42	–0.42
	—	—	(0.66)	(0.66)	(0.66)	(0.67)	(0.68)	(0.68)	(0.68)
Out-to-In	—	—	–0.07**	–0.12***	–0.12***	–0.06**	–0.08**	–0.07*	–0.07*
	—	—	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)
Generation	—	—	—	0.06	0.10	0.18	0.18	0.19	0.18
	—	—	—	(0.09)	(0.16)	(0.18)	(0.18)	(0.19)	(0.19)
Subcomponents	—	—	—	—	–0.001	0.04***	0.04***	0.03***	0.03***
	—	—	—	—	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Assets	—	—	—	—	—	–0.034	–0.04	–0.04	–0.04
	—	—	—	—	—	(0.03)	(0.04)	(0.04)	(0.05)
Employees	—	—	—	—	—	—	0.01	0.01	0.01
	—	—	—	—	—	—	(0.06)	(0.07)	(0.06)
Patents	—	—	—	—	—	—	—	0.01	0.01
	—	—	—	—	—	—	—	(0.05)	(0.05)
Volume	—	—	—	—	—	—	—	—	0.01
	—	—	—	—	—	—	—	—	(0.05)
R^2	0.31	0.32	0.33	0.37	0.38	0.38	0.38	0.38	0.38
N	2.54	2.54	2.54	2.54	2.54	2.54	2.54	2.54	2.54

Notes. This table presents estimates of our defect-rate model using our secondary identification technique. We present the estimates in the fashion of Table 8. The sample horizon is 1999–2014. We calculate standard errors with the bootstrap, sampling supply chain links randomly.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Table 13. Quality Improvement Model Estimates

	1	2	3	4	5	6	7	8	9
Distance	0.21***	0.18***	0.20***	0.21***	0.21***	0.22***	0.23***	0.22***	0.22***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Luxury	—	–0.08	–0.01	–0.09	–0.08	–0.08	–0.08	–0.09	–0.09
	—	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.015)
International	—	—	–0.03	–0.02	–0.04	–0.05	–0.062	–0.06	–0.06
	—	—	(0.22)	(0.22)	(0.22)	(0.23)	(0.23)	(0.23)	(0.23)
Generation	—	—	—	0.11*	0.12	0.12	0.12	0.13*	0.13
	—	—	—	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)
Subcomponents	—	—	—	—	0.10***	0.10***	0.103***	0.10***	0.10***
	—	—	—	—	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Assets	—	—	—	—	—	0.04***	0.05***	0.05**	0.04***
	—	—	—	—	—	(0.01)	(0.01)	(0.01)	(0.01)
Employees	—	—	—	—	—	—	–0.01	–0.02	–0.02
	—	—	—	—	—	—	(0.05)	(0.05)	(0.05)
Patents	—	—	—	—	—		—	0.02	0.02
	—	—	—	—	—		—	(0.04)	(0.04)
Volume	—	—	—	—	—		—	—	0.02***
	—	—	—	—	—		—	—	(0.00)
Constant	–1.56***	–1.29***	–1.494***	–2.334***	–2.686***	–3.2503***	–3.263***	–3.155***	–3.313***
	(0.42)	(0.16)	(0.43)	(0.74)	(0.78)	(0.85)	(0.88)	(0.88)	(1.12)
R^2	0.07	0.09	0.12	0.14	0.16	0.16	0.17	0.17	0.18
N	1.78	1.78	1.78	1.78	1.78	1.78	1.78	1.78	1.78

Notes. This table presents estimates of our defect-rate model by estimating the evolution of defect rates for a given product, produced by the same supplier for the same car model across subsequent model iterations. We present the estimates in the fashion of Table 8. The sample horizon is 1999–2014.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Figure 8. (Color online) Geographic Distances vs. Travel Distances

Notes. This figure shows a visualization of the geographic distance vs. travel distance for five suppliers of the Ford Explorer 2013, which is assembled in Chicago, Illinois. The supplier factories are located in Fayetteville, Arkansas (tires); Pensacola, Florida (braking system); Carey, Ohio (seating); Livonia, Michigan (doors/tailgate); and Ladson, South Carolina (axles). For each route, we obtained the respective travel time by looking at average traffic congestion.

it only has one factory. Our results are virtually identical whether we include suppliers with multiple factories or restrict our sample to single-factory suppliers.

10.2. Competition

The degree of supplier competition is a potential confounding variable. The competitiveness of the environment near the assembly plant can causally affect a component's quality—a supplier will work harder when the local competition is fierce—and how far the assembler must go to source it. An automaker, after all, won't have to go as far when the market is thick. In this fashion, supplier competitiveness could mediate the effect of supplier proximity on product quality.

To measure this effect, we add three additional independent variables: the number of suppliers of the given component within 100,500 and 1,000 km of the given assembly plant. Table 15 shows evidence of a competition effect: defect rates decrease significantly with all three supplier counts. But we find that our results are robust—still at the 1% level—with the inclusion of the distance variables.

10.3. Cluster Robust Standard Errors

We change our standard error methodology from the bootstrap to a two-way cluster-robust approach. These standard errors permit general error term heteroskedasticity and select error term cross-correlations. Specifically, two errors may be correlated if they have a "cluster" in common; we cluster our sample by supplier and assembly plant, so $\varepsilon_{p_1s_1}$ may covary with $\varepsilon_{p_2s_2}$ if $s_1 = s_2$ or if products p_1 and p_2 are installed in cars at the same facility. For example, Ford's Hermosillo assembly plant corresponds to 421 observations, which may have a general error term covariance matrix, and supplier Yorozu Corp. corresponds to 65 observations, which may also have a general error term covariance matrix. The effect of supply chain distance is also statistically significant with these alternative standard errors.

10.4. B-Spline Terms

To quantify the benefit of including the distance-based information in our supplier-selection model, we compare our model to an alternative without the B-splines. We run the regressions at the part-type level. For the average part, our supplier-selection model has a root-mean-square error (RMSE) of 0.1723 with the distance variables and an RMSE of 0.1898 without the distance variables. Thus, our RMSE would on average be (0.1898–0.1723)/0.1723 = 8.42% larger without these. Further, to verify if our results hold without the B-splines, we run the regressions without the selection probability B-spline terms. Removing the Dahl correction terms does not meaningfully change our results.

10.5. Probit Selection Probabilities

To test whether our results are robust to the form of the supplier-selection probability, we change our selection probability estimates from multinomial logistic regression fitted values to standard probit regression fitted values (with the same regressors). We find that changing the form of the supplier-selection probability does not meaningfully change our results.

10.6. Additive Estimates

In the base model, we use a multiplicative (log-log) model to account for the positive and skewed nature of the continuous variables. We now change our econometric specification from multiplicative to additive, removing logarithms. We also rescale the continuous variables to have a standard deviation of one, to make their coefficients more interpretable. We estimate that increasing supply chain distance by one standard deviation increases the expected failure rate by a significantly positive amount.

10.7. Paired Fixed Effects

Since close coordination and collaboration regarding component designs is a key conjectured mechanism for attaining proximity benefits, we are likely to observe different outcomes for different country pairs. To ensure that our results are not affected by this peculiarity, we obtain estimates for a model that pairs the upstream-location × downstream-location fixed effects. Column 7 in Table 15 shows that our estimates are robust to this estimation.

10.8. Internal Combustion vs. Electric/ Hybrid Vehicles

Electric cars have lower component complexity and higher software complexity than internal combustion (IC) cars. Thus, the strength of the relationship between

Table 14. Defect-Rate Model Estimates Under Travel Time Distance Measure

	1	2	3	4	5	6	7	8	9
Travel Time	0.95**	0.99**	1.03***	0.97**	0.93**	1.00***	0.99**	1.03***	1.03***
	(0.40)	(0.40)	(0.39)	(0.38)	(0.37)	(0.38)	(0.38)	(0.39)	(0.39)
Luxury	—	3.42***	3.42***	3.56***	3.15***	3.22***	3.21***	3.20***	3.37***
	—	(0.62)	(0.62)	(0.82)	(0.80)	(0.81)	(0.82)	(0.81)	(0.83)
International	—	—	–0.27	–0.40	–0.46	–0.54	–0.52	–0.56	–0.47
	—	—	(0.80)	(0.79)	(0.76)	(0.76)	(0.76)	(0.75)	(0.74)
Generation	—	—	—	–1.71**	–2.20***	–2.33***	–2.33***	–2.34***	–3.02***
	—	—	—	(0.79)	(0.80)	(0.81)	(0.80)	(0.81)	(0.85)
Subcomponents	—	—	—	—	3.61***	3.60***	3.60***	3.60***	2.73***
	—	—	—	—	(0.35)	(0.36)	(0.35)	(0.35)	(0.43)
Assets	—	—	—	—	—	0.61***	0.63***	0.67***	0.68***
	—	—	—	—	—	(0.18)	(0.17)	(0.18)	(0.17)
Employees	—	—	—	—	—	—	–0.07	–0.05	–0.02
	—	—	—	—	—	—	(0.14)	(0.14)	(0.14)
Patents	—	—	—	—	—	—	—	–0.12	–0.09
	—	—	—	—	—	—	—	(0.16)	(0.16)
Volume	—	—	—	—	—	—	—	—	12.82***
	—	—	—	—	—	—	—	—	(2.61)
Dahl coefficient 1	–3.97	–3.91	–3.90	–3.89	–3.90	–3.98	–3.92	–3.92	–3.97
	(3.42)	(3.41)	(3.41)	(3.41)	(3.34)	(3.32)	(3.36)	(3.38)	(3.35)
Dahl coefficient 2	–4.18*	–4.21*	–4.21*	–4.00*	–3.95*	–4.06*	–4.02*	–3.94*	–4.20*
	(2.21)	(2.21)	(2.20)	(2.18)	(2.16)	(2.17)	(2.20)	(2.24)	(2.21)
Dahl coefficient 3	–5.09**	–5.05**	–5.05**	–5.11**	–5.08**	–5.27**	–5.16**	–4.94**	–4.87**
	(2.15)	(2.16)	(2.16)	(2.15)	(2.12)	(2.13)	(2.21)	(2.32)	(2.27)
Dahl coefficient 4	2.61	2.59	2.59	3.45	3.45	3.55	3.61	3.84	3.33
	(4.81)	(4.79)	(4.79)	(4.77)	(4.72)	(4.70)	(4.71)	(4.81)	(4.73)
Dahl coefficient 5	–13.88*	–13.61*	–13.61*	–15.30*	–14.65*	–14.95*	–14.79*	–14.73*	–13.96*
	(8.02)	(7.97)	(8.00)	(8.24)	(8.11)	(8.19)	(8.24)	(8.21)	(7.93)
Dahl coefficient 6	1.94	1.76	1.90	3.83	1.78	1.03	1.20	1.78	0.78
	(10.15)	(10.07)	(10.07)	(10.42)	(10.48)	(10.57)	(10.62)	(10.79)	(10.31)
R^2	0.21	0.21	0.21	0.21	0.22	0.22	0.22	0.22	0.22
VIF	—	1.00	1.14	1.17	1.17	1.27	1.55	1.87	1.90
N	24.07	24.07	24.07	24.07	24.07	24.07	24.07	24.07	24.07

Notes. This table presents estimates of the defect-rate coefficients in the fashion of Table 8. For these regressions, however, we measure *Distance* with the expected travel time, rather than with geographic distance, calculated with Vicenty's formula. The sample horizon is 1999–2014.

$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Distance and *Defect Rate* may change as the fleet transitions away from IC vehicles. We test whether this is the case by dividing our sample into parts that are specific to IC vehicles (such as engines, transmissions, and fuel systems) and parts that are not specific to IC vehicles (such as tires, seats, and airbags). Table 17 in the online Appendix indicates that the non-IC-specific parts represent the majority of our sample with respect to both counts and defect reports. We rerun our primary analysis on the subsample of parts that are not specific to internal combustion vehicles and tabulate the regression coefficients in column 9 of Table 15, finding that even when we rule out the complexities due to internal combustion, there is still a strong relationship between *Distance* and *Defect Rate*. This result suggests that removing the IC engine won't be enough to overcome the burden of distance in auto supply chains.

10.9. Truncated Dependent Variable

We rerun our analysis by estimating the complaint rate with only the defects in year x, for $x \in \{1, 2, 3\}$. These estimates are presented in columns 9, 10, and 11 in Table 15. The coefficient estimates of *Distance* are larger with only these early year defects. This makes sense because early year defects are more likely to stem from production and design problems (which should depend on supply chain distance), whereas latter-year defects are more likely to arise from normal wear and tear (which shouldn't depend on supply chain distance).

10.10. Nonlinear Impact of Distance

Figure 9 depicts the relationship between distance and quality nonparametrically along the percentiles of the *Distance* distribution. The relationship appears linear,

Table 15. Robustness Checks

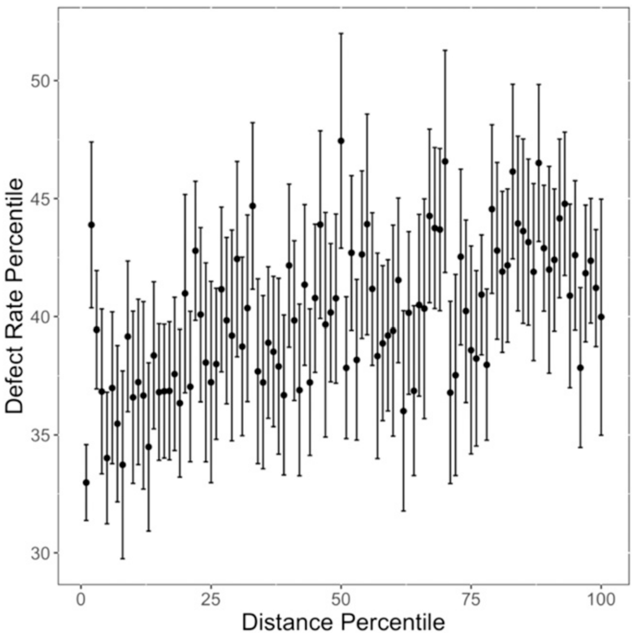
	1	2	3	4	5	6	7	8	9	10	11	12
Distance	0.82***	0.55***	0.88***	0.81***	0.81***	2.03***	0.47***	0.62***	1.62***	1.40***	1.36**	−2.63**
	(0.17)	(0.15)	(0.17)	(0.16)	(0.16)	(0.68)	(0.23)	(0.18)	(0.47)	(0.51)	(0.55)	(1.24)
Distance^2												0.23***
												(0.09)
Luxury	2.93***	2.86***	2.88***	2.85***	2.85***	21.43***	2.76***	4.66***	−8.05***	−5.13***	−6.39***	2.98***
	(0.59)	(0.58)	(0.59)	(0.57)	(0.57)	(2.79)	(0.76)	(0.94)	(2.42)	(1.69)	(1.89)	(0.72)
International	0.41	0.37	−0.12	0.04	0.04	1.67	0.38	−0.13	−3.83**	−1.90	0.31	0.17
	(0.63)	(0.60)	(0.66)	(0.62)	(0.63)	(2.12)	(0.99)	(0.60)	(1.88)	(1.67)	(2.10)	(0.57)
Generation	−1.22***	−1.22***	−0.89	−4.33***	−1.41***	−1.39***	−1.25***	−2.14***	−1.11***	−3.97***	−6.96***	−1.38***
	(0.31)	(0.31)	(0.65)	(0.36)	(0.32)	(0.34)	(0.42)	(0.49)	(0.21)	(1.18)	(1.76)	(0.44)
Subcomponents	3.72***	3.92***	3.58***	3.77***	3.77***	20.99***	3.71***	2.85***	4.71***	4.75***	4.43***	3.75***
	(0.20)	(0.20)	(0.19)	(0.19)	(0.19)	(0.91)	(0.25)	(0.24)	(0.88)	(0.79)	(0.71)	(0.24)
Volume	20.54***	20.55***	14.54***	9.63***	20.46***	20.23***	21.07***	17.41***	47.11***	28.88***	38.78***	20.81***
	(1.42)	(1.42)	(2.61)	(0.80)	(1.31)	(1.46)	(1.11)	(1.38)	(4.85)	(4.58)	(4.03)	(1.12)
Assets	0.09	0.10	0.08	0.08	0.08	1.37	0.07	−0.07	0.13	0.13	−0.46	0.11
	(0.11)	(0.11)	(0.12)	(0.11)	(0.11)	(2.01)	(0.14)	(0.07)	(0.30)	(0.33)	(0.35)	(0.13)
Employees	−0.13	−0.09	−0.01	−0.07	−0.07	−1.26	−0.09	−0.06	−0.29*	−0.23	0.18	−0.09
	(0.10)	(0.10)	(0.11)	(0.10)	(0.10)	(1.57)	(0.09)	(0.11)	(0.17)	(0.16)	(0.21)	(0.09)
Patents	−0.32***	−0.19**	−0.19*	−0.19**	−0.19**	−3.19*	−0.19**	−0.02	−1.00***	0.03	−0.02	−0.20**
	(0.10)	(0.09)	(0.10)	(0.09)	(0.09)	(1.77)	(0.08)	(0.12)	(0.30)	(0.37)	(0.42)	(0.09)
Dahl coefficient 1	1.71	−0.31	0.41	—	−16.76	2.04	2.11	−3.74	2.85	−14.76	−19.53	1.99
	(1.82)	(1.72)	(1.68)	—	(40.50)	(3.85)	(5.45)	(4.40)	(60.45)	(49.29)	(48.61)	(5.61)
Dahl coefficient 2	−1.64	−0.94	−1.33	—	5.18	−0.85	−6.87**	−2.42	31.41	8.64	18.77	−6.08*
	(1.17)	(1.21)	(1.18)	—	(13.83)	(4.02)	(3.38)	(3.44)	(22.96)	(25.65)	(42.49)	(3.35)
Dahl coefficient 3	1.00	−0.67	−0.05	—	−1.55	1.59	−0.09	0.21	10.65	−3.45	1.11	0.04
	(1.30)	(1.36)	(1.31)	—	(18.36)	(4.04)	(1.97)	(1.66)	(26.41)	(21.52)	(26.38)	(1.92)
Dahl coefficient 4	1.42	1.29	1.15	—	−11.39	6.60	−0.39	5.17	12.75	−2.23	19.25	−0.13
	(2.23)	(2.27)	(2.14)	—	(17.78)	(9.49)	(3.71)	(4.30)	(60.73)	(62.48)	(53.72)	(3.73)
Dahl coefficient 5	−0.51	−4.77	−3.22	—	17.84	−12.84	9.43	1.72	34.97	−15.81	−25.41	9.81
	(3.36)	(3.56)	(3.39)	—	(28.16)	(14.42)	(6.31)	(6.83)	(72.85)	(144.63)	(89.24)	(6.51)
Dahl coefficient 6	−3.02	−2.28	−2.95	—	−12.13	−4.27	−16.35*	−6.56	−7.28	−2.70	10.16	−16.56*
	(6.04)	(5.62)	(6.01)	—	(21.16)	(22.53)	(8.44)	(7.18)	(24.55)	(30.90)	(38.94)	(8.59)
Suppliers(100)	—	−1.45***	—	—	—	—	—	—	—	—	—	—
	—	(0.51)	—	—	—	—	—	—	—	—	—	—
Suppliers(500)	—	−0.82*	—	—	—	—	—	—	—	—	—	—
	—	(0.46)	—	—	—	—	—	—	—	—	—	—
Suppliers(1000)	—	−2.80***	—	—	—	—	—	—	—	—	—	—
	—	(0.77)	—	—	—	—	—	—	—	—	—	—
R^2	0.25	0.25	0.28	0.21	0.26	0.25	0.26	0.16	0.32	0.29	0.29	0.25
VIF	1.92	1.92	1.91	5.76	2.03	1.99	1.92	2.00	1.90	1.91	1.95	106.72
N	23.79	23.79	23.79	23.79	21.79	22.29	23.70	16.11	23.79	19.11	17.70	23.79

Notes. This table presents coefficients of the defect-rate model under seven alternate specifications. The column 1 regression limits the sample to contracts won by suppliers with only one factory. The column 2 specification adds three new variables to defect rate model: the number of supplier factories that produce the given part within 100, 500, and 1,000 km of the assembly plant. The column 3 specification uses the cluster-robust standard errors of Cameron et al. (2008). The column 4 specification has no correction terms; that is, we eliminate the Dahl coefficients. The column 5 specification estimates supplier-selection probabilities with a probit regression, rather than a multinomial logistic regression. The column 6 specification is additive, with all logarithms removed from the independent and dependent variables (for formatting purposes, we also scale each variable so its standard deviation is one). The column 7 specification adds fixed effects for every assembler country, supplier-country pair. The column 8 specification only estimates the impact on non-IC-specific components. Columns 9–11 present estimates of our defect rate model with only the reports received in the initial x years, where $x = \{1, 2, 3\}$. Column 12 presents estimates of our defect rate model with a quadratic fit. The sample horizon is 1999–2014.

$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

which justifies our linear statistical specifications. To study this assumption more carefully, we run a quadratic fit for the *Distance* variable. Column 12 in Table 15 reports coefficient estimates from a model that permits *Defect Rate* to depend on both *Distance* and *Distance*^{2}. Overall, 97% of our sample lies in the region in which *Defect Rate* increases with *Distance* (i.e., the region of the parabola with positive slope). So even with a quadratic model, we find positive association between distance and defect rates.

Figure 9. Mean and Median Effect across the Distribution of Distance



Notes. To create this plot, we grouped the data into 100 buckets, based on *Distance* percentiles and normalized by brand and part. We then calculated each bucket's mean brand-part-specific *Defect Rate* percentile and its corresponding 90% confidence interval.

11. Conclusion

Supply chain proximity facilitates coordination, which facilitates continuous improvement. Dyer (1996, pp. 282–287) hypothesized

> a relationship between quality and the adjusted measure of face-to-face contact . . . . Toyota's value chain is more specialized, and more productive, than Nissan's or GM's, and a key factor seems to be the geographic proximity of Toyota's production network. Indeed, we might expect a geographically concentrated production network to have advantages because proximity facilitates the formal and informal dissemination of information and technology across firms (as evidence, our survey found that Toyota and its suppliers engaged in the most information sharing). Moreover, proximity makes available a range of human-intensive (high media) communication mechanisms which facilitates the flow of tacit and complex knowledge across firms (Daft and Lengel, 1986, Almeida and Kogut, 1994). Thus, Toyota's network may simply learn faster than Nissan's or GM's network.

We provide the first empirical evidence of Dyer's conjecture: we find it is more difficult to extract a reliable part from a far-off land. We estimate that replacing the current supplier distance with that of the next-closest competitor would increase the average defect rate by 1.57%–2.17%.

> "So, what would a one-percent change [in the defect rate] mean?" Nick Persichilli, a Senior Manager at the Automotive Part's Manufacturer's Association (APMA), asks. "A one-percent change would mean the world, because we're always looking to improve efficiencies, find improvements in quality, find better ways to manufacture it. . . . There is a lot of interest, and it would mean a lot to the industry." In addition to Mr. Persichilli, we asked three NHTSA executives to interpret our results. These interviews impressed upon us the fact that the brunt of poor auto quality is borne not by the industry, but by the public at large. And this public expense is formidable: in 2015, NHTSA estimated the economic and societal cost of auto collisions to be $242 and $847 billion per year, respectively. In the Online Appendix, we rerun our analysis with NHTSA's crash reports, rather than defect reports, as the dependent variable. We estimate that doubling all supply chain distances would increase crash rates by 0.37%, which would increase the societal cost of auto collisions by 0.37% * $847 billion per year = $2.4 billion per year.[4]

Guajardo et al. (2016, p. 1872) estimate that "for a car with median product characteristics in [their] sample, the effect on consumer utility of a 1% price decrease is equivalent, all else being equal, to increasing product quality by 2.2%." Combining this result with ours, we estimate that replacing the current set of suppliers with the next-closest set yields a quality change roughly equal to a 1% price change.

Acknowledgments

The authors thank Jacob Dubé for his excellent research assistance and Hana Galal for her early feedback on this manuscript.

Endnotes

[1] We use Vincenty's (1975) formula to calculate geographic distances from latitudes and longitudes.

[2] We present supplier-selection coefficients from the first stage in the online Appendix.

[3] Both measures—*Geographic Distance* and *Travel Time*—have advantages. On the one hand, measuring proximity with *Travel Time* yields a more direct proxy of the ease of collaboration, monitoring, and face-to-face interactions (Giroud 2013). On the other hand, *Geographic Distance* can be defined without imposing assumptions on the firms' travel behavior, whereas travel time forces us to make assumptions about the modes of transportation, travel routes, congestion levels, etc.

[4] See the online Appendix for additional analysis based on auto recalls.

References

Adams B (2013) Migration and agglomeration among motor vehicle parts suppliers. Working paper, California State University, East Bay, Hayward.

Almeida P, Kogut B (1999) Technology and geography: The localization of knowledge and the mobility of patent holders. *Management Sci*. 45(7):905–1024.

Berking J, Birchler B, Buss J, Jade L, Stolz L, Kern D (2016) Insights on automotive supplier excellence: Footprint optimization at automotive suppliers. Technical report, Oliver Wyman, New York. Accessed January 1, 2018, https://www.oliverwyman.com/content/dam/oliver-wyman/v2/publications/2016/jan/Supplier_Footprint_Final_Web.pdf.

Bernard AB, Moxnes A, Saito YU (2015) Production networks, geography and firm performance. Technical report, National Bureau of Economic Research, Cambridge, MA.

Bodenheimer G (2014) Mitigating packaging damage in the supply chain. *Packaging Digest* (September 11), http://www.packagingdigest.com/supply-_chain/mitigating-_packaging-_damage-_inthe-_supply-_chain140910.

Brülhart M, Trionfetti F (2001) Industrial specialisation and public procurement: Theory and empirical evidence. *J. Econom. Integr.* 16(1):106–127.

Cachon GP, Olivares M (2010) Drivers of finished-goods inventory in the US automobile industry. *Management Sci.* 56(1):202–216.

Cameron AC, Gelbach JB, Miller DL (2008) Bootstrap-based improvements for inference with clustered errors. *Rev. Econom. and Statist*. 90(3):414–427.

Clark KB, Fujimoto T (1991) *Product Development Performance: Strategy, Organization, and Management in the World Auto Industry* (Harvard Business School Press, Boston).

Colak A, Bray RL (2016) Why do automakers initiate recalls? A structural econometric study. Working paper, Clemson University, Clemson, SC.

Dahl GB (2002) Mobility and the return to education: Testing a Roy model with multiple markets. *Econometrica* 70(6):2367–2420.

Dawson C (2010) Toyota revives dormant plant. *Wall Street Journal* (December 27), https://www.wsj.com/articles/SB10001424052970204527804576043681007542352.

Dyer JH (1996) Specialized supplier networks as a source of competitive advantage: Evidence from the auto industry. *Strategic Management J.* 17(4):271–291.

Dyer JH (2000) *Collaborative Advantage: Winning Through Extended Enterprise Supplier Networks* (Oxford University Press, New York).

Eaton D (2015) Honda moving Acura MDFX production to Ohio. *Columbus Bus. First* (February 3), https://www.bizjournals.com/columbus/news/2015/02/11/honda-moving-acura-mdx-production-to-ohio.html.

Gawer A (2011) *Platforms, Markets and Innovation* (Edward Elgar Publishing, Cheltenham, UK).

Giroud X (2013) Proximity and investment: Evidence from plant-level data. *Quart. J. Econom*. 128(2): 861–915.

Guajardo JA, Cohen MA, Netessine S (2016) Service competition and product quality in the U.S. automobile industry. *Management Sci.* 62(7):1860–1877.

Hendricks KB, Singhal VR (2005) An empirical analysis of the effect of supply chain disruptions on long-run stock price performance and equity risk of the firm. *Production Oper. Management* 14(1):35–52.

Jain N, Girotra K, Netessine S (2014) Managing global sourcing: Inventory performance. *Management Sci.* 60(5):1202–1222.

Guajardo JA, Cohen MA, Kim S-H, Netessine S (2012) Impact of performance-based contracting on product reliability: An empirical analysis. *Management Sci.* 58(5):961–979.

Newey WK (1984) A method of moments interpretation of sequential estimators. *Econom. Lett.* 14(2-3):201–206.

Osadchiy N, Gaur V, Seshadri S (2016) Systematic risk in supply chain networks. *Management Sci.* 62(6):1755–1777.

Pete J (2015) Ford to invest $900 million, hire 200 more at Chicago plant. *NWI Times* (November 11), http://www.nwitimes.com/business/local/ford-_to-_invest-_million-_hire-_more-_at-_chicago-_plant/article_20ac0e97-_4667-_5b14-_a416-_464ea00b66a4.html.

Ramdas K, Randall T (2008) Does component sharing help or hurt reliability? An empirical study in the automotive industry. *Management Sci.* 54(5):922–938.

Rosenbaum T (2014) Where do automotive suppliers locate and why? Working paper, Yale University, New Haven, CT.

Salerno MS, Zilbovicius M, Arbix G, Carneiro Dias AV (1998) Changes and persistences on the relationship between assemblers and suppliers in Brazil. *Actes du GERPISA* 24:51–66.

Schmitt A, Van Biesebroeck J (2013) Proximity strategies in outsourcing relations: The role of geographical, cultural and relational proximity in the European automotive industry. *J. Internat. Bus. Stud.* 44(5):475–503.

Serpa JC, Krishnan H (2017) The impact of supply chains on firm-level productivity. *Management Sci.* 64(2):511–532.

Shah R, Ball GP, Netessine S (2016) Plant operations and product recalls in the automotive industry: an empirical investigation. *Management Sci.* 63(8):2439–2459.

Shingal A (2015) Econometric analyses of home bias in government procurement. *Rev. Internat. Econom.* 23(1):188–219.

Spindelndreier D, Bauernhansl T, Beck M, Lesmeister F (2015) The proximity paradox: Balancing auto suppliers' manufacturing networks. Report, The Boston Consulting Group, Boston.

Trionfetti F (2000) Discriminatory public procurement and international trade. *World Econom.* 23(1):57–76.

Tunca TI, Zhu W (2018) Buyer intermediation in supplier finance. *Management Sci.* 64(12):5631–5650.

Udenio M, Gaur V, Fransoo JC (2014) Inventories and the credit crisis: A chicken and egg situation. Working paper, Eindhoven University of Technology, Eindhoven, Netherlands.

Vincenty T (1975) Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations. *Survey Rev*. 23(176):88–93.

Wang YI, Li J, Anupindi R (2017) Risky suppliers or risky supply chains? An empirical analysis of sub-tier supply network structure on firm performance in the high-tech sector. Ross School of Business Paper No. 1297, University of Michigan, Ann Arbor.

Woodruff D, Templeman J (1993) Why Mercedes is Alabama bound. Accessed April 9, 2016, https://www.bloomberg.com/news/articles/1993-10-10/why-mercedes-is-alabama-bound.

Wu Di (2016) Shock spillover and financial response in supply chain networks: Evidence from firm-level data. Working paper, University of Michigan, Ann Arbor.