

NBER WORKING PAPER SERIES

INNOVATION IN THE GLOBAL FIRM

L. Kamran Bilir
Eduardo Morales

Working Paper 22160
<http://www.nber.org/papers/w22160>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2016

We thank Costas Arkolakis, Paula Bustos, Jan De Loecker, Amit Gandhi, Elhanan Helpman, Sam Kortum, Nina Pavcnik, Veronica Rappoport, Joel Rodrigue, Andrés Rodríguez-Clare, Felix Tintelnot, Jim Tybout, David Weinstein, Steve Yeaple, Bill Zeile, and seminar participants at AEA, CEPR, Colorado-Boulder, Columbia, Dartmouth, Duke, George Washington University, HBS, Hebrew University, LMU, LSE, Michigan, Minnesota, NBER, Kellogg, Penn State, Princeton, Santa Cruz, SED, UCLA, UPenn, Vanderbilt, and Wisconsin for very helpful comments. The statistical analysis of firm-level data on U.S. multinational companies was conducted at the Bureau of Economic Analysis, U.S. Department of Commerce, under arrangements that maintain legal confidentiality requirements. The views expressed are those of the authors and do not reflect official positions of the U.S. Department of Commerce or the National Bureau of Economic Research. The Bureau of Economic Analysis has reviewed this paper prior to its release to ensure that data confidentiality is not unintentionally compromised. Bilir thanks the International Economics Section at Princeton and the Wisconsin Alumni Research Foundation for support. Morales thanks the University of Wisconsin-Madison and the Cowles Foundation at Yale University for their hospitality and support.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by L. Kamran Bilir and Eduardo Morales. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Innovation in the Global Firm
L. Kamran Bilir and Eduardo Morales
NBER Working Paper No. 22160
April 2016, Revised November 2016
JEL No. F00,F23,O30

ABSTRACT

How global are the gains from innovation? When firms operate production plants in multiple countries, technological improvements developed in one location may be shared with foreign sites for efficiency gain. We develop a model that accounts for such transfer, and apply it to measure private returns to R&D investment for a panel of U.S. multinationals during 1989{2008. Our estimates indicate that innovation increases performance at firm locations beyond the innovating site: the average U.S. multinational firm realizes abroad over a quarter of the return to its U.S. R&D investment, suggesting estimates based only on domestic operations understate multinationals' gain from innovation. We also find that multinationals' U.S. R&D significantly increases aggregate value added in foreign countries that host their affiliates, revealing a spatial disconnect between the costs and potential gains of policies that encourage firms' U.S. innovation.

L. Kamran Bilir
Department of Economics
University of Wisconsin - Madison
1180 Observatory Drive
Madison, WI 53706
kbilir@ssc.wisc.edu

Eduardo Morales
Department of Economics
Princeton University
310 Fisher Hall
Princeton, NJ 08544
and NBER
ecmorales@princeton.edu

1 Introduction

Multinational corporations are among the most innovation intensive firms and account for the majority of innovation investment worldwide.¹ Although defined by their fragmentation of production across countries, innovation within these firms is spatially concentrated by comparison, with a large share of firms pursuing innovation investment in only one (headquarters) country.² If this concentrated investment results in technological improvements that are shared with foreign sites for efficiency gain, these facts raise the strong possibility that location-specific policies encouraging innovation create within-firm gains that are realized abroad. While this is a stated concern of the U.S. Congress, which awards over \$30 billion in innovation subsidies to U.S. firms annually (National Science Board 2014), its importance hinges on the actual extent to which innovation within the multinational firm affects the productivity of affiliates abroad.³

This paper quantifies the *intrafirm* impact of U.S. innovation investment by multinational firms on the performance dynamics of manufacturing affiliates abroad. Guided by a model of global innovation and production, we estimate this impact of innovation for a panel of U.S.-based firms with affiliates operating in 48 countries during 1989–2008. Our results provide evidence that headquarters innovation significantly increases affiliate performance within the same multinational firm, and is also the primary determinant of long-run affiliate performance. Quantitatively, we find that the average firm thus realizes *abroad* over a quarter of the return to its U.S. R&D investment, suggesting estimates based only on domestic operations understate firms’ gain from U.S. innovation. Aggregating across firms, our estimates further suggest that multinationals’ U.S. innovation substantially raises aggregate value added in foreign countries hosting their affiliates. Affiliate innovation, by contrast, does not affect performance growth at other firm sites.

Estimating the impact of headquarters innovation on affiliate performance requires taking a stand on the production structure of multinational firms. Guided by recent evidence indicating intrafirm input trade is low within U.S. multiplant firms (Atalay, Hortaçsu, and Syverson 2014) and U.S. multinational firms (Ramondo, Rappoport, and Ruhl 2016), our main specification features ‘horizontal’ firms that do not engage in intrafirm goods trade but instead sell output to arm’s-length markets. Within this structure, affiliates produce output subject to an idiosyncratic performance level that evolves according to a Markov process. This performance process may respond to R&D investments of the affiliate itself, its U.S. headquarters (parent), or other foreign affiliates within the same firm, reflecting the possible effects of innovation on affiliate performance growth. To evaluate the performance impact of each innovation source, we build on the contributions of Aw, Roberts, and Xu (2011) and Doraszelski and Jaumandreu (2013). Our approach is novel in its explicit consideration of the headquarters innovation impact on foreign affiliate performance within

¹UNCTAD (2005), Criscuolo, Haskel, and Slaughter (2010). In the case of the United States, multinationals account for 91 percent of the innovation investment by firms (National Science Board 2014).

²For U.S. multinationals, see Table 1 below.

³Large corporations claim the majority of U.S. R&D tax credits (Government Accountability Office 2009). These corporations can indefinitely defer paying U.S. taxes for profits earned by foreign affiliates. The U.S. Senate has thus stated that: “Substantial amounts of income [are] earned [...] not in the United States where I think most of it probably belongs. [...] Right now it seems clear to me we are not getting our fair share. The R&D is done here. It is supported with an R&D tax credit. [...] This is not in balance,” U.S. Senate (2013).

the same firm—an impact that is important not only for the spatial disconnect it reveals between the costs and potential gains of U.S. innovation policy, but also for what it reflects regarding the importance of intangible input transfer across firm sites. In particular, our estimates suggest that ownership theories emphasizing the intrafirm transfer of technological inputs are relevant for explaining the existence of multinational firms.⁴

To estimate the parameters of the model, we use detailed affiliate-level panel data on the global operations of U.S.-based multinational firms from the Bureau of Economic Analysis (BEA). These data allow us to observe separate measures of parent- and affiliate-specific R&D spending, as well as inputs and output, within each multinational firm and year. The availability of data on the allocation of innovation investment across sites within the same firm is particularly unique and important; such information is rarely available to researchers, but is essential in our analysis. Our primary results consider firms in the computer industry; we show that our main qualitative findings also hold in the pharmaceutical and motor vehicles industries, as well as three broader associated sectors: industrial machinery, chemicals, and transportation equipment. These are the three largest manufacturing sectors among U.S. multinational firms (U.S. BEA 2008). We observe that affiliates within a firm often produce distinct products and therefore base our main estimates on the complete set of foreign manufacturing affiliates in which the parent owns a majority stake; we also provide separate estimates based only on affiliates operating in the parent industry.

The estimates we recover indicate that the performance of an affiliate is persistent and increasing in its own R&D investment. Our estimates also reveal that affiliate performance increases in the innovation of its U.S. parent, and that parent and affiliate R&D are complementary. This latter result adds nuance to the standard view of the multinational firm (Helpman 1984, Markusen 1984), which considers all affiliates as pure recipients rather than producers of technology. It is nevertheless consistent with Cohen and Levinthal (1989), which proposes that innovation enhances a firm’s ability to assimilate and exploit information—including, in this case, information resulting from R&D performed at the parent site. We find that the performance impact of an affiliate’s own innovation investment is significant, but strictly secondary to that of parent R&D. Furthermore, omitting parent innovation as a determinant of affiliate performance leads to an overstatement—of approximately 50 percent—in the performance impact of affiliates’ own innovation investment.

Our estimates reveal that the firm-wide, value-added return to U.S. parent R&D is understated considerably when only its local impact on the firm’s U.S. operations is accounted for. Specifically, we find that the gross private return to headquarters R&D investment, defined as the impact of an infinitesimal increase in parent R&D on the total value added earned by its multinational firm, exceeds the parent-level return by 37 percent for the average multinational firm. Our estimates also imply that parent innovation is a critical determinant of long-run affiliate performance: eliminating the effect of parent R&D would, all else equal, imply an average reduction in affiliate performance of 62 percent; by contrast, eliminating an affiliate’s own R&D would result in an analogous decline of just 18 percent. This impact of parent innovation on long-run affiliate performance further translates into an effect on aggregate affiliate value added, thus implying that countries hosting

⁴For example, Arrow (1975), Teece (1982), Atalay, Hortaçsu, and Syverson (2014).

high levels of U.S. affiliate activity would observe a significant decline in industry-level GDP if affiliates were suddenly unable to benefit from U.S. parent innovation—an effect that may be further magnified by technology spillovers from affiliates to domestic firms (Javorcik 2004a) and links with domestic input producers (Rodríguez-Clare 1996). Importantly, such technological knowledge spillovers are known to be substantially local (Keller 2002); this, combined with our estimates and the dominance of U.S. parent innovation in aggregate U.S. R&D spending, supports the idea that multinational production is an important determinant of technology diffusion between countries.⁵

To identify the performance impact of innovation, we consider specifications that project measures of current affiliate performance on different lags of parent and affiliate R&D investment. Importantly, these specifications control for country-industry-year fixed effects and corresponding lags of affiliate performance. Consistent with our assumption that affiliate performance follows a Markov process, controlling for lagged performance ensures that our estimates are not affected by reverse causality. Specifically, our estimates do not reflect the spurious correlation between lagged R&D and current performance that would otherwise arise if performance is both persistent and correlated with contemporaneous R&D spending. In addition, the country-industry-year fixed effects we include account for a range of time-varying, unobserved factors that may simultaneously affect both affiliate performance growth and affiliate incentives to perform R&D. For example, countries with an abundance of skilled labor may be attractive locations for affiliate R&D investment, and affiliates in such countries may also have higher performance levels.

A potential concern that remains is that other determinants of future affiliate performance may be known to the firm at the time innovation decisions take place, but not controlled for by our measures of lagged affiliate performance and country-industry-year effects. For example, affiliate performance could be affected by persistent exogenous factors that are known to the firm, but that affect affiliates differentially even within the same country-industry market. This consideration is unlikely to be quantitatively important when estimating the impact of parent R&D on affiliate performance within large multinational firms, however: parent R&D decisions likely depend on the set of such unobserved affiliate-specific factors across *all* firm sites, and are therefore less sensitive to the idiosyncratic performance shocks of an individual foreign affiliate. This reasoning is consistent with the data; the average multinational firm has 14 foreign affiliates that together account for only approximately 25 percent of total sales and employment, so that each affiliate tends to be small compared with the overall firm.

A second potential concern is that multinationals may misreport innovation investment for tax purposes: firms may, for example, intentionally overreport R&D spending by affiliates located in high-tax countries with the aim of underreporting profits. To account for this possibility, we build instruments for affiliate R&D, interacting the user cost of R&D corresponding to the U.S. parent location (U.S. state) from Wilson (2009) with an index of intellectual property rights protection in the affiliate host country from Ginarte and Park (1997) and Park (2008). We find that, even after controlling for country-year and firm fixed effects, observed affiliate R&D spending is significantly higher in countries with strong intellectual property rights when the corresponding U.S. parent

⁵This has long been hypothesized in the literature; Keller (2004) provides a detailed review.

also faces high R&D costs. Importantly, the country-industry-year fixed effects we include in our estimation ensures that this instrument is valid even if intellectual property rights are correlated with local policies that affect incentives to misreport, such as corporate tax rates.⁶ Notice, however, that this instrumentation strategy cannot solve concerns like that described in the paragraph above. This is because a serially correlated unobservable renders endogeneous not only R&D, but also all production inputs and their respective lagged values.

This paper contributes to a literature evaluating the impact of R&D investment on plant-level and domestic firm-level outcomes, and methodologically follows Aw, Roberts, and Xu (2011) and Doraszelski and Jaumandreu (2013); our focus on global firms further relates our analysis to Bøler, Moxnes, and Ulltveit-Moe (2015). We contribute to this research by estimating the impact of parent R&D investment on the performance of foreign manufacturing affiliates within the multinational firm, a magnitude that gains relevance in light of multinationals' dominance in worldwide innovation spending. Given that the multinational firm context we consider thus involves not only multiple countries, but also dynamics and interdependence across firm locations, we pursue an estimation approach closely resembling Doraszelski and Jaumandreu (2013); that is, we leave unspecified the innovation cost function of the firm and estimate the impact of innovation on performance without requiring a full solution to the firm's complex dynamic problem.

The estimates we recover support theories of the firm featuring sites linked by intangible transfers, and thus complement existing evidence including Atalay, Hortaçsu, and Syverson (2014). They find evidence consistent with the importance of intrafirm intangible transfers by showing that, after a change in U.S. plant ownership, acquired establishments evolve to resemble the acquiring firm despite a lack of physical shipments linking sites within the multiplant firm.⁷ While this approach takes a general view of the form of the intangibles involved, our analysis provides related evidence for a specific form of intangible input that we find is transferred across sites within the firm: the proprietary knowledge resulting from R&D investment. A key motivation for our focus on R&D is the prevalence, in many countries, of policies that subsidize local firms' innovation, and the corresponding potential for these subsidies to impact multinational firms' performance abroad.

Our results are closely related to work aiming to establish the existence of international technology transfer across plants within the multinational firm (Branstetter, Fisman, and Foley 2006; Keller and Yeaple 2013; Gumpert 2015).⁸ A distinction in our paper is that we infer the flow of technology by estimating the impact of U.S. parent R&D on foreign affiliate performance, without relying on observed proxies for technology transfer. We are therefore able to compare the conclusions that arise from these two approaches for measuring technology transfer within the multinational firm. Consistent with Branstetter, Fisman, and Foley (2006), we find that the payment of

⁶We also address potential R&D reporting errors in specifications that exclude affiliates located in tax havens.

⁷Ramondo, Rappoport, and Ruhl (2016) document a similar lack of shipments across sites within U.S. multinational firms, and Arnold and Javorcik (2009), Guadalupe, Kuzmina, and Thomas (2012), and Javorcik and Poelhekke (2016) show that affiliates acquired by foreign multinationals are faster growing than unaffiliated firms.

⁸Branstetter, Fisman, and Foley (2006) use royalties paid to parents by affiliates as a proxy for technology transfer. Keller and Yeaple (2013) use the joint distribution of U.S. multinationals' intrafirm trade and affiliate sales to show that firm activity is consistent with parents and affiliates sharing technology in both tangible and intangible forms. Gumpert (2015) presents empirical evidence for knowledge transfers using data on corporate transferees. Giraud (2013) provides indirect evidence of intrafirm trade in intangibles within U.S. multiplant firms.

royalties and technology license fees is indeed positively correlated with the affiliate-level impact of parent innovation that we estimate. In addition, consistent with Keller and Yeaple (2013), we extend our baseline framework to a ‘vertical’ specification that permits parent-affiliate trade, and find that intrafirm goods trade channels proprietary technology across firm sites.

Finally, the quantitative implications of intrafirm technology transfer that we compute are related to studies assessing implications of input trade and proprietary technology transfer within multinational firms, including McGrattan and Prescott (2010), McGrattan (2012), Irrazabal, Moxnes, and Opromolla (2013), and Bilir (2014). Our estimates complement research investigating the welfare gains from multinational production and the importance of the cross-plant, within-firm productivity distribution for the magnitude of these gains (Ramondo and Rodríguez-Clare 2013, Arkolakis et al 2014, Head and Mayer 2016, Tintelnot 2016)—a distribution that in our model depends on endogenous decisions of the firm, including the extent of parent and affiliate R&D investments.

The rest of the paper is organized as follows. Section 2 presents an empirical model of innovation in the multinational firm. Section 3 describes the data. Section 4 outlines the baseline estimation strategy and discusses our identification assumptions. Sections 5 and 6 present our estimates, and section 7 discusses quantitative implications of these estimates. Sections 8 and 9 extend the baseline model to account for intrafirm trade and misreporting, respectively, and section 10 concludes. Derivations and additional details may be found in the online Appendix.

2 Empirical Model

This section describes an empirical model of production and innovation investment in the multinational firm. The baseline model considers a ‘horizontal’ firm that does not engage in intrafirm trade across plants, consistent with evidence in Ramondo, Rappoport, and Ruhl (2016).⁹ We use our model to derive estimating equations that may be combined with available data to determine the performance impact of innovation across different sites within the multinational firm.

2.1 Setup

Time is discrete. Consider a set of multinational firms $i = 1, \dots, \mathcal{I}_t$ operating within the same manufacturing industry, defined as the output industry of the parent. The set of firm- i production sites active in period t is \mathcal{J}_{it} . Sites in \mathcal{J}_{it} are indexed by j , where $j = 0$ denotes the parent and $j > 0$ corresponds to its foreign manufacturing affiliates. We focus in this section on foreign affiliates, and postpone our separate treatment of parents to section 7.2.

2.2 Demand

Within firm i , each affiliate j sells a single variety as a monopolistically competitive firm in a market n_{ij} . We define n_{ij} as the country-sector pair in which affiliate j produces, and assume that

⁹Ramondo, Rappoport, and Ruhl (2016) report that the majority of foreign affiliates of U.S. multinationals send no shipments to, and receive below one percent of inputs from, their U.S. parent. Section 8 and Appendix A.14 nevertheless consider an extended version of our model in which affiliates may source inputs from their parent.

any two distinct affiliates j and j' of firm i operate in distinct markets n_{ij} and $n_{ij'}$.¹⁰ Thus, while each manufacturing affiliate operates in a single sector, different affiliates within a given firm may produce in different industries and in different countries.¹¹ Importantly, the country corresponding to n_{ij} is the firm- i , affiliate- j production location, but need not be the location of its customers.¹²

Assume that affiliate j faces the following demand function for its output Q_{ijt} :

$$Q_{ijt} = Q_{n_{ijt}}(P_{ijt}/P_{n_{ijt}})^{-\sigma} \exp[\xi_{ijt}(\sigma - 1)], \quad (1)$$

where $\sigma > 1$ is the elasticity of substitution across output varieties, P_{ijt} is the output price set by affiliate j , and ξ_{ijt} is an unobserved demand shock (or product quality shock) that is known to the firm when making its input, output, and pricing decisions at period t .¹³ Market-level variables $P_{n_{ijt}}$ and $Q_{n_{ijt}}$ denote the period- t aggregate price index and demand level, respectively.

2.3 Production

To produce output Q_{ijt} , affiliate j combines capital, labor, and materials using the following production technology

$$Q_{ijt} = (H(K_{ijt}, L_{ijt}; \alpha))^{1-\alpha_m} M_{ijt}^{\alpha_m} \exp(\omega_{ijt}) \quad (2)$$

where

$$H(K_{ijt}, L_{ijt}; \alpha) = \exp(h(k_{ijt}, l_{ijt}; \alpha)) \quad (3)$$

$$h(k_{ijt}, l_{ijt}; \alpha) \equiv \alpha_l l_{ijt} + \alpha_k k_{ijt} + \alpha_{ll} l_{ijt}^2 + \alpha_{kk} k_{ijt}^2 + \alpha_{lk} l_{ijt} k_{ijt} \quad (4)$$

and $\alpha = (\alpha_l, \alpha_k, \alpha_{ll}, \alpha_{kk}, \alpha_{lk})$.¹⁴ In (2), K_{ijt} is effective units of capital, L_{ijt} is the number of production workers, M_{ijt} is an unobserved quantity index of materials use, and ω_{ijt} denotes the Hicks-neutral physical productivity at t .¹⁵ Consistent with our baseline assumption that firm i

¹⁰This definition is in line with the data, which do not distinguish between plants owned by firm i that operate in the same country-industry pair. This market definition accommodates a range of firm structures, from those with affiliates operating in the same sector in spatially segmented markets, to those with affiliates operating in different sectors in the same geographic market.

¹¹Both assumptions are consistent with the data. In 2004, for example, over 60 percent of U.S. computer firms' foreign affiliates produce exclusively in one three-digit industry. Among the approximately 40 percent remaining, the affiliate's primary industry accounts for 92 percent of its total sales, on average. However, within the same firm, affiliates specialize in a range of industries: for example, only 16 percent produce in the same industry as their U.S. parent. Affiliates are similarly heterogeneous in their geographic location. The United Kingdom, which hosted the largest number of U.S. firm affiliates in 2004, accounted for just above 10 percent of affiliates in that year.

¹²Our estimation approach (see section 4) does not require taking a stand on the geographic location of demand, and thus implicitly accommodates export-platform sales.

¹³This demand function has previously been used in the production function literature (e.g. De Loecker 2011). Other work emphasizes the empirical relevance of both heterogeneous and variable markups (De Loecker and Warzynski 2012, De Loecker et al 2016), including in contexts with endogenous innovation (Doraszelski and Jaumandreu 2013). Markup heterogeneity may be accommodated in our framework through heterogeneity in σ . Section 6 and Appendix A.5 extend our model to allow σ to vary by year. Allowing for variable markups given the demand in (1) requires relaxing the assumed monopolistic competition structure; see Appendix A.6 for further discussion.

¹⁴Lower-case Latin letters denote the logarithm of the upper-case variable, e.g. $l_{ijt} = \ln(L_{ijt})$.

¹⁵A recent literature estimates production functions that include both Hicks-neutral and labor-augmenting productivity terms; see Bøler (2015), Doraszelski and Jaumandreu (2016), for example. We leave the application of this

does not engage in intrafirm input trade, we assume here that M_{ijt} includes only inputs purchased from arms'-length suppliers. This production function in (2) combines materials with a translog function of capital and labor, defined in (3) and (4), according to a Cobb-Douglas technology.¹⁶ The elasticity of output with respect to materials is captured in (2) by α_m ; output elasticities with respect to capital and labor may be heterogeneous across affiliates in (4), reflecting differences in factor usage. We assume affiliates take prices of labor P_{ijt}^l , capital P_{ijt}^k , and materials P_{ijt}^m as given, and that the latter is common to all affiliates within a market-year: $P_{ijt}^m = P_{n_{ijt}}^m$.¹⁷

2.4 Value Added Function

Given the production and demand structures described above, and assuming firm i determines M_{ijt} optimally by maximizing affiliate- j static profits at t , log value added va_{ijt}^* may be expressed as

$$va_{ijt}^* = \kappa_{n_{ijt}} + h(k_{ijt}, l_{ijt}; \beta) + \psi_{ijt}, \quad (5)$$

where $\beta = \alpha(1 - \alpha_m)\iota$ and $\iota = (\sigma - 1)/(\sigma - \alpha_m(\sigma - 1))$; see Appendix A.1 for a derivation of this expression. In (5), $\kappa_{n_{ijt}}$ is a function of the materials price $P_{n_{ijt}}^m$, aggregate price index $P_{n_{ijt}}$, and aggregate demand level $Q_{n_{ijt}}$ in market n_{ij} at t ; $h(\cdot)$ is the translog function of capital and labor inputs in (4) above. The term ψ_{ijt} is the ι -scaled sum of the affiliate- j physical productivity and demand (product quality) shock: $\psi_{ijt} \equiv \iota(\omega_{ijt} + \xi_{ijt})$. We refer to ψ_{ijt} as the idiosyncratic performance level of affiliate j .¹⁸

Allowing value added to be measured with error, we express observed value added $va_{ijt} \equiv va_{ijt}^* + \varepsilon_{ijt}$ as

$$va_{ijt} = \kappa_{n_{ijt}} + h(k_{ijt}, l_{ijt}; \beta) + \psi_{ijt} + \varepsilon_{ijt}, \quad (6)$$

and assume that the measurement error ε_{ijt} is mean independent of capital k_{ijt} , labor l_{ijt} , and all variables known to firm i in period $t - 1$:

$$\mathbb{E}_{t-1}[\varepsilon_{ijt} | k_{ijt}, l_{ijt}] = 0. \quad (7)$$

2.5 Impact of Innovation on Firm Performance

The performance ψ_{ijt} of firm i 's affiliate j evolves over time according to the stochastic process

extension to the context of multinational firms for future work.

¹⁶The elasticity of substitution between materials and the joint output of capital and labor is restricted to one by necessity: affiliate materials use is not directly observed in the data. Nevertheless, (2) yields a value added function analogous to that in the literature (e.g. De Loecker and Warzynski 2012). While we could attempt to recover a proxy for spending on materials as the difference between revenue and value added, any measurement error in these latter variables would also impact this proxy, impeding the estimation of a production function that is translog in materials. See Appendix A.7 for discussion.

¹⁷Assuming affiliates in the same market-year share a common material input price enables us to account for it through market-year effects (see sections 2.4 and 4).

¹⁸Performance thus combines supply and demand shifters as in Foster, Haltiwanger and Syverson (2008), De Loecker (2011), and Bøler, Moxnes, and Ulltveit-Moe (2015).

$$\psi_{ijt} = \mathbb{E}_{t-1}[\psi_{ijt}] + \eta_{ijt}, \quad (8)$$

where, in our baseline setting, the expectation of affiliate- j performance ψ_{ijt} conditional on the information of firm i at $t - 1$ is

$$\mathbb{E}_{t-1}[\psi_{ijt}] = \rho\psi_{ijt-1} + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \mu_{n_{ijt}}. \quad (9)$$

This expectation depends on a) lagged affiliate- j performance ψ_{ijt-1} through the persistence parameter ρ , b) lagged affiliate- j R&D investment r_{ijt-1} and parent R&D investment r_{i0t-1} through the parameters μ_a , μ_p and μ_{ap} , and c) a market-year unobserved term $\mu_{n_{ijt}}$ reflecting country-sector-year characteristics that affect the performance of all affiliates operating in n_{ij} at t .¹⁹ We define $r_{ijt} = 0$ for observations in which affiliate j does not perform R&D.²⁰ From (8), η_{ijt} captures exogenous shocks affecting the performance of j at t that are not anticipated by i at $t - 1$. To account for the possibility that the distribution of these performance shocks may differ across firms, countries, and years, our estimation approach does not restrict the distribution of η_{ijt} beyond the mean-independence condition implied by (8). Specifically, in section 6 we present results with standard errors clustered at the firm-year level, thereby allowing η_{ijt} to reflect firm-year specific shocks that affect simultaneously all sites j within firm i .

The expected productivity of affiliate j in (9) depends on its past productivity and R&D spending, in line with recent plant-level models of innovation including Aw, Roberts, and Xu (2011), Doraszelski and Jaumandreu (2013), and Bøler, Moxnes, and Ulltveit-Moe (2015).²¹ A distinction in (9) is the inclusion of R&D investment performed by the parent of affiliate j . This allows us to assess the intrafirm influence of headquarters innovation on the performance of foreign affiliates. Specifically, positive values of μ_p and μ_{ap} in (9) would be consistent with parent innovation increasing affiliate performance, with an impact that is itself increasing in affiliates' own R&D spending. Sections 5 and 6 present parameter estimates corresponding to the baseline model in (9) and several alternative specifications. Section 8 includes a generalization of (9) that permits affiliate imports from the parent to influence affiliate performance.²²

¹⁹Equations (8) and (9) describe the evolution of performance ψ_{ijt} . Identifying separate processes for ω_{ijt} and ξ_{ijt} would require observing either output prices (see Roberts et al 2011) or revenue in at least two separate markets per affiliate (see Jaumandreu and Yin 2014); neither is available in our dataset. Key parameters μ_a , μ_p and μ_{ap} may thus be interpreted as reflecting the joint impact of R&D on ω_{ijt} (process innovation) and ξ_{ijt} (product innovation). Appendix A.2 contains examples of different stochastic process for productivity ω_{ijt} and demand shifters ξ_{ijt} that are consistent with the evolution of performance described by equations (8) and (9).

²⁰In the estimation sample, 92.5 percent of parent observations have a positive R&D expenditure, while just 24.1 percent of affiliate observations do (see Table 1). We measure R_{ijt} in thousands of U.S. dollars. Among non-zero observations for affiliate R&D, all satisfy $R_{ijt} > 1$, and r_{ijt} is therefore strictly positive among R&D-performing sites. In section 6.2, we relax the functional form in equation (9), estimating separate market-year fixed effects for affiliates performing R&D and for those with $R_{ijt} = 0$.

²¹In our context, introducing nonlinear functions of ψ_{ijt-1} in (9) poses an empirical challenge. As section 4.1 shows, our estimation approach requires including a large set of market-year fixed effects. These would enter the estimating equation nonlinearly if we were to allow for higher-order terms such as ψ_{ijt-1}^2 in (9); due to the resulting incidental parameters problem, our parameter estimates would be asymptotically biased. See Appendix A.8 for details.

²²A control that has been shown in the literature to be important (e.g. Evans 1987, Dunne et al 1989, Huergo and Jaumandreu 2004, Haltiwanger, Jarmin, and Miranda 2013) and that we are not able to include in (9) is affiliate age.

2.6 Firm Optimization

In every period t , firm i determines optimal levels of labor \mathbf{L}_{it} , materials \mathbf{M}_{it} , capital investment \mathbf{I}_{it} , R&D investment \mathbf{R}_{it} , and output prices \mathbf{P}_{it} for each of its affiliates active at t , and also determines the set of affiliates that will be active at $t + 1$, \mathcal{J}_{it+1} .²³ These decisions are a function of firm i 's state vector \mathbf{S}_{it} , which has elements

$$S_{ijt} = (\psi_{ijt}, K_{ijt}, P_{ijt}^l, P_{ijt}^k, P_{n_{ijt}}^m, Q_{n_{ijt}}, P_{n_{ijt}}, \mu_{n_{ijt}}, \chi_{ijt}^k, \chi_{ijt}^r, F_{ijt}), \quad (10)$$

where χ_{ijt}^k and χ_{ijt}^r are exogenous affiliate-specific shocks to the cost of investment in physical capital and R&D, respectively, and F_{ijt} is a fixed operating cost.

The Bellman equation associated with firm i 's dynamic optimization problem is

$$V(\mathbf{S}_{it}) = \max_{\mathbf{C}_{it}} \left\{ \sum_{j \in \mathcal{J}_{it}} \Pi(S_{ijt}, I_{ijt}, L_{ijt}, M_{ijt}, P_{ijt}, R_{ijt}) + \delta \mathbb{E}[V(\mathbf{S}_{it+1}) | \mathbf{S}_{it}, \mathbf{I}_{it}, \mathbf{R}_{it}] \right\} \quad (11)$$

where $\mathbf{C}_{it} = \{\mathcal{J}_{it+1}, \mathbf{I}_{it}, \mathbf{L}_{it}, \mathbf{M}_{it}, \mathbf{P}_{it}, \mathbf{R}_{it}\}$ is the set of control variables, $V(\cdot)$ is the value function, $\Pi(\cdot)$ is the profit function, and δ is the discount factor. We assume capital at t is determined by physical capital investment in all previous periods according to the law of motion $K_{ijt} = \delta_k K_{ijt-1} + I_{ijt-1}$. If active at period t , the profit function of firm i 's affiliate j is

$$\Pi(S_{ijt}, I_{ijt}, L_{ijt}, M_{ijt}, P_{ijt}, R_{ijt}) = VA_{ijt}^* - W_{ijt}^l - C_k(P_{ijt}^k, I_{ijt}, K_{ijt}, \chi_{ijt}^k) - C_r(R_{ijt}, \chi_{ijt}^r) - F_{ijt} \quad (12)$$

where $C_k(\cdot)$ and $C_r(\cdot)$ are cost functions of investment in physical capital and R&D, and $W_{ijt}^l \equiv P_{ijt}^l L_{ijt}$ is total spending on labor inputs. Our estimation approach below does not require specifying $C_k(\cdot)$, $C_r(\cdot)$, or the distributions of cost shocks χ_{ijt}^k , χ_{ijt}^r , and F_{ijt} .²⁴

2.7 Interpreting Innovation Effects in the Model

The parameters μ_a , μ_p and μ_{ap} in (9) jointly capture the impact of innovation investments \mathbf{R}_{it-1} on performance ψ_{ijt} for manufacturing affiliates present at both $t - 1$ and t . These innovation

Our data do not provide information on the age of affiliates resulting from the acquisition of a previously existing firm (which, according to the BEA data, form around 65% of all manufacturing affiliates).

²³The baseline model thus assumes affiliate entry and exit decisions are taken with a one-period lag; section 4.2 considers instead the case of instantaneous affiliate entry and exit. For any X , we henceforth denote $\mathbf{X}_{it} = \{X_{ijt}\}_{j \in \mathcal{J}_{it}}$.

²⁴The absence of restrictions on $C_k(\cdot)$, $C_r(\cdot)$ and the distributions of χ_{ijt}^k and χ_{ijt}^r ensures that there exist $C_k(\cdot)$ and $C_r(\cdot)$ functions as well as χ_{ijt}^k and χ_{ijt}^r realizations that are able to rationalize any observed pattern of investment in R&D and physical capital, including observed zeros. Specifically, these cost shocks allow for differences in both fixed and variable R&D and capital investment costs. Estimating $C_k(\cdot)$ and $C_r(\cdot)$ using necessary conditions for optimality of observed R&D and capital investment requires accounting for interdependence in these decisions across affiliates and over time, and therefore solving a dynamic discrete choice problem with a very large choice set and state vector. This poses a well-known computational challenge (e.g. Holmes 2011, Morales, Sheu, and Zahler 2015), and resulting estimates and predictions may be very sensitive to the definition of firms' information sets (Dickstein and Morales 2015). Separately, Fillat and Garetto (2015), Gumpert et al (2016), and Garetto, Oldenski, and Ramondo (2016) have shown that a sunk cost of affiliate entry is able to explain important features of multinational firm dynamics. While for notational simplicity, the profit function in equation (12) does not include such a sunk cost, the estimation approach, results, and implications described below would not be affected by including such costs.

parameters reflect a range of potential channels through which innovation may impact affiliate performance. These include the impact of R&D investment as manifested through affiliates that upgrade product quality, increase manufacturing efficiency, or switch to a new product.

In addition, firms in our model choose whether to a) become a multinational by opening a first affiliate, b) maintain an existing affiliate, or c) open a new affiliate based, in part, on the performance of all potential firm sites. Our model thus accounts for the possibility that innovation impacts affiliate entry and exit through its effect on affiliate performance. Although μ_a , μ_p and μ_{ap} do not themselves quantify the affiliate entry effects of innovation, the estimation framework accounts for—and is thus consistent with—responsiveness to innovation along these entry margins.²⁵

The functional form in (9) is motivated by two considerations. First, the model nests a benchmark specification, studied recently in Aw, Roberts and Xu (2011) and Doraszelski and Jaumandreu (2013), whereby the productivity of a domestic firm or plant responds exclusively to its own innovation investment—that is, innovation that is performed within the same domestic firm or plant. This specification is captured by a version of (9) that restricts both μ_p and μ_{ap} to be zero. Estimated values of these parameters significantly different from zero would thus suggest the importance of a new, international dimension of intrafirm R&D effects on the evolution of plant performance within the multinational firm.²⁶ Second, regarding the interaction term governed by μ_{ap} , notice that in its absence, the log-linear functional form implies a specific degree of substitutability between parent and affiliate R&D in determining subsequent affiliate performance. By including this interaction term, we relax this restriction, allowing μ_{ap} to flexibly govern the degree of substitutability between affiliate and parent R&D as determinants of affiliate performance.

Finally, the model treats μ_a , μ_p and μ_{ap} as technological parameters. An alternative would be a model in which these parameters reflect the combination of a) knowledge communication frictions, and b) the decision over *how much* proprietary technology a parent transmits to its foreign affiliate. Given the data in hand, however, we are unable to distinguish between these two forces.²⁷

3 Data and Measurement

Estimating the parameters of the model in section 2 requires measures of innovation for each multinational firm parent, and measures of inputs, output, and innovation for each of its affiliates. We describe these below.

3.1 Innovation and Production in U.S. Multinational Firms

We use affiliate-level panel data on the global operations of U.S.-based multinational firms from the Bureau of Economic Analysis (BEA) Survey of U.S. Direct Investment Abroad. These confidential

²⁵Our model abstracts from the potential markup effects of innovation, though extended versions could accommodate this possibility. See Appendix A.6 for discussion.

²⁶An alternative would be to regress the performance of each affiliate on its own R&D investment and on the total R&D expenditure of its multinational firm. The results from such a specification are likely to be similar to those obtained from (9) given the dominance of parent R&D in the overall R&D investment of each multinational firm.

²⁷For recent work that considers the firm decision over technology transfer to affiliates within the multinational firm, see Bilir (2014) and Holmes, McGrattan, and Prescott (2015).

data provide information on U.S. parent companies and each foreign affiliate on an annual basis.²⁸ For our analysis, we assemble separate datasets corresponding to different manufacturing industries for the period 1989–2008. Details regarding data construction appear in Appendix A.3.

The data include direct affiliate-level measures that correspond to variables in the model. These include value added VA_{ijt} , the value of physical plant, property, and equipment, net of depreciation K_{ijt} , the number of employees L_{ijt} , and total employee compensation W_{ijt}^l . Affiliate-level measures of inputs sourced from the parent and technology license fees paid to the parent are also available and will enter our analysis in sections 6.1 and 8, respectively. The index of materials use M_{ijt} , the output market price and quantity indexes $P_{n_{ijt}}$ and $Q_{n_{ijt}}$, and the materials price $P_{n_{ijt}}^m$ are not observed. Importantly, the data include separate parent- and affiliate-level measures of R&D investment for distinct sites within the same firm. The availability of panel data that include measures of production inputs, output, and R&D by site within the same firm is both essential for estimating the model in section 2 and also unusual. To our knowledge, the data provided by the BEA is the only affiliate-level resource that provides a homogeneous measure of site-level innovation spending within a comprehensive panel of multinational or multiplant firms.

The measure of innovation investment in the data captures primarily variable costs of performing R&D. Specifically, it includes spending on wages and salaries, materials, and supplies used in both basic and applied R&D, and also spans the range between product and process R&D. This definition is consistent with the model in section 2, which considers R&D investment as impacting performance, which itself reflects both production efficiency ω_{ijt} and product quality ξ_{ijt} .²⁹ This measure does not account for spending on capital inputs, routine testing and quality control, market research, or legal expenses related to patents.

Labor L_{ijt} in the model is a production input, but a plant performing innovation may dedicate a subset of its labor to R&D activity. The data do not always include separate measures of production and innovation employees, and our baseline estimation therefore measures L_{ijt} as the (always available) total number of employees. Benchmark-year surveys do, however, record separate measures of both total employment and R&D employment. Using these, we construct an alternative measure of L_{ijt} that, similar to Schankerman (1981), corrects for affiliate-specific differences in the share of total employment devoted to innovation; we use this measure in section 6.3.³⁰

We estimate the parameters of the model separately by industry. Multinationals are assigned to an industry based on that of the parent. Our main analysis uses only information on manufacturing

²⁸The survey is conducted by the BEA for the purpose of producing publicly available statistics on the operations of U.S. multinationals and is comprehensive in its coverage. Any U.S. person having direct or indirect ownership or control of 10 percent or more of the voting securities of an incorporated foreign business enterprise or an equivalent interest in an unincorporated foreign business enterprise at any time during the survey fiscal year in question is considered to have a foreign affiliate. The country of an affiliate corresponds to the location of its physical assets.

²⁹This aspect of the data precludes estimating separate effects for product and process innovation; see Cohen and Klepper (1996) and Dhingra (2013) for models that feature both types of innovation. Comprehensive benchmark-year surveys in 1989, 1994, 1999, and 2004 provide a decomposition of R&D spending according to the entity paying for the R&D and the entity performing the R&D, which may in certain cases be distinct. In line with the model assumptions, this decomposition shows that nearly all of the R&D activity completed at an affiliate site is also paid for by the performing affiliate (U.S. BEA 2008).

³⁰The data do not provide information to permit an analogous approach for capital inputs. However, we find in section 6.3 that correcting measured labor inputs has only a negligible impact on the estimates obtained, suggesting this is unlikely to be an important empirical concern.

affiliates; our estimation sample thus excludes affiliates in agriculture, mining, construction, transportation and public utilities, finance, insurance, retail, real estate services, health services, and other services. In section 6.4, we demonstrate the robustness of our results to further restricting our sample to include only affiliates operating in the parent industry.

The main analysis evaluates firms operating in Computers and Office Equipment (SIC 357), a manufacturing industry that accounts for the production of electronic computers, computer terminals, computer peripheral equipment, calculating and accounting machines, and office machines. Summary statistics for this industry appear in Table 2. To assess the sensitivity of our results to potential differences across industries, section 6.5 presents estimates for multinational firms in Motor Vehicles and Equipment (SIC 371) and Pharmaceutical Drugs (SIC 283). We further evaluate the robustness of our results to the degree of sectoral aggregation in the data; section 6.5 provides estimates that rely on a broader industry definition: Chemicals (SIC 28), Industrial Machinery (SIC 35), and Transportation Equipment (SIC 37). These are the three largest manufacturing sectors according to the number of affiliates (U.S. BEA 2008).

3.2 Descriptive Statistics

Our main estimation sample for the Computers and Office Equipment industry includes 119 parent firms located in 29 different U.S. states, and their 942 majority-owned, foreign manufacturing affiliates located in 48 countries. The modal affiliate also operates in the computer industry (SIC 357), and the average parent firm has 14 foreign affiliates (standard deviation 10.4).

Table 1 summarizes the distribution of innovation investment within and across U.S.-based multinational firms in the computer industry. The top rows reveal three salient features of the data regarding the organization of innovation activity across sites within the firm: first, nearly all U.S. parents invest in R&D; second, only 47.8 percent of multinational firms have at least one foreign affiliate performing R&D; and third, for the average firm, fewer than 25 percent of affiliates perform any R&D. These statistics indicate that the spatial concentration of innovation investment is higher than that of production within multinational firms: although multinationals are defined by their fragmentation of production across countries, the majority of these firms perform R&D investment in only one country. In Appendix A.3, we provide analogous statistics for the three broader industries considered in our estimation, Chemicals, Industrial Machinery, and Transportation Equipment, and find that the within-firm organization of innovation in each of these industries is similar to that in the computer industry.

Another way to understand the dispersion of production relative to innovation is to consider the share of each activity accounted for by foreign affiliates within the firm. Rows 4 through 8 in Table 1 show that affiliates in the computer industry tend to account for only a small share—8.5 percent, on average—of firm-level R&D spending. By contrast, the affiliate share in total employment for the average firm is 26.6 percent, and in sales is 28.8 percent. The affiliate share in value added, 16.7 percent, is also higher than that in R&D spending. A similar pattern holds in each of the broader industries we consider (Table A.1, Appendix A.3). Parent sites are thus responsible for a substantially higher share of innovation investment than of production, on average. That the spatial

concentration of innovation investment is higher than that of production is consistent with the idea that knowledge is shared across firm locations (Arrow 1975, Teece 1982). However, the observation that some firms fragment innovation across countries may suggest the presence of frictions limiting the communication of technical knowledge across sites (Arrow 1962, 1969). Importantly, these statistics suggest that a subset of affiliates are producers rather than pure recipients of knowledge.

4 Estimation

This section uses the model described in section 2 to derive a set of estimating equations. We show how these may be combined with the data in section 3 to estimate the parameters of the model determining the affiliate-level performance impacts of parent and affiliate innovation.

4.1 Estimation Approach

The parameters of interest are those determining affiliate value added in (6), β , and those governing the short-run and long-run impacts of innovation on affiliate performance in (9), $(\rho, \mu_a, \mu_p, \mu_{ap})$. To derive an estimating equation, we first combine (6), (8), and (9), to arrive at

$$\begin{aligned} va_{ijt} = & h(k_{ijt}, l_{ijt}; \beta) + \rho(va_{ijt-1} - h(k_{ijt-1}, l_{ijt-1}; \beta)) \\ & + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \gamma_{n_{ijt}} + u_{ijt}, \end{aligned} \quad (13)$$

where $u_{ijt} \equiv \eta_{ijt} + \varepsilon_{ijt} - \rho\varepsilon_{ijt-1}$ is a function of the performance shock and measurement error in value added, and where $\gamma_{n_{ijt}} \equiv \mu_{n_{ijt}} + \kappa_{n_{ijt}} - \rho\kappa_{n_{ijt-1}}$ is a market-year effect that accounts for both the unobserved quantity and prices embedded in $\kappa_{n_{ijt}}$ and the market-year specific component of firm performance, $\mu_{n_{ijt}}$.

Estimating (13) requires addressing two identification challenges. First, as a static (flexible) input, labor hired by firm i 's affiliate j during period t is determined after the period- t shock to performance η_{ijt} is observed by firm i . This gives rise to a correlation between l_{ijt} and the error term u_{ijt} (Griliches and Mairesse 1998). Second, u_{ijt} is also a function of the measurement error ε_{ijt-1} in value added va_{ijt-1} , resulting in a correlation between va_{ijt-1} and u_{ijt} . To simultaneously address both challenges, we estimate the parameters of interest in two steps.

The first step uses the affiliate labor optimality condition to estimate parameters determining the elasticity of value added with respect to labor $(\beta_l, \beta_{ll}, \beta_{lk})$ and the measurement error component of value added ε_{ijt} for each affiliate and period (see Gandhi, Navarro, and Rivers 2016). Conditional on these first-step estimates, the second step estimates the remaining model parameters $(\beta_k, \beta_{kk}, \rho, \mu_a, \mu_p, \mu_{ap})$ while controlling for the market-year effects $\{\gamma_{n_{ijt}}\}$. These steps are described below.

Step 1 Given the production function in (2), the profit function in (12), and the assumption that labor is a static input, a necessary condition for labor to be optimally determined is

$$\beta_l + 2\beta_{ll}l_{ijt} + \beta_{lk}k_{ijt} = \frac{W_{ijt}^l}{VA_{ijt}} \exp(\varepsilon_{ijt}), \quad (14)$$

where, as in (12), W_{ijt}^l is total affiliate- j spending on labor inputs during period t . Parameters $(\beta_l, \beta_{ll}, \beta_{lk})$ are thus identified from the moment condition in (7), which implies

$$\mathbb{E}[va_{ijt} - w_{ijt}^l + \log(\beta_l + 2\beta_{ll}l_{ijt} + \beta_{lk}k_{ijt}) | l_{ijt}, k_{ijt}, j \in \mathcal{J}_{it}] = 0, \quad (15)$$

where conditioning on $j \in \mathcal{J}_{it}$ reflects that identification relies only on affiliates active at period t . Given (15), we use Nonlinear Least Squares (NLS) to estimate $(\beta_l, \beta_{ll}, \beta_{lk})$.³¹ With the estimates $(\hat{\beta}_l, \hat{\beta}_{ll}, \hat{\beta}_{lk})$ in hand, we recover an estimate of the measurement error ε_{ijt} for each firm i , affiliate j , and period t : $\hat{\varepsilon}_{ijt} = va_{ijt} - w_{ijt}^l + \log(\hat{\beta}_l + 2\hat{\beta}_{ll}l_{ijt} + \hat{\beta}_{lk}k_{ijt})$.³²

Step 2 Using the estimates $(\hat{\beta}_l, \hat{\beta}_{ll}, \hat{\beta}_{lk})$ and $\hat{\varepsilon}_{ijt}$, we construct $\widehat{va}_{ijt} \equiv va_{ijt} - \hat{\beta}_l l_{ijt} - \hat{\beta}_{ll} l_{ijt}^2 - \hat{\beta}_{lk} l_{ijt} k_{ijt} - \hat{\varepsilon}_{ijt}$ and rewrite (13) as

$$\begin{aligned} \widehat{va}_{ijt} = & \beta_k k_{ijt} + \beta_{kk} k_{ijt}^2 + \rho(\widehat{va}_{ijt-1} - \beta_k k_{ijt-1} - \beta_{kk} k_{ijt-1}^2) \\ & + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \gamma_{n_{ijt}} + \eta_{ijt}. \end{aligned} \quad (16)$$

Notice that the error term in (16) is now simply the performance shock, η_{ijt} . We thus base the identification of $(\beta_k, \beta_{kk}, \rho, \mu_a, \mu_p, \mu_{ap})$ on (8), which implies

$$\mathbb{E}[\eta_{ijt} | k_{ijt}, \widehat{va}_{ijt-1}, k_{ijt-1}, r_{ijt-1}, r_{i0t-1}, \gamma_{n_{ijt}}, j \in \mathcal{J}_{it-1}, j \in \mathcal{J}_{it}] = 0. \quad (17)$$

Notice that, by conditioning on both $j \in \mathcal{J}_{it-1}$ and $j \in \mathcal{J}_{it}$, (17) accounts for the restriction that an affiliate must be active in both periods $t-1$ and t to be included in the estimation sample. Using (16) and (17), we estimate $(\beta_k, \beta_{kk}, \rho, \mu_a, \mu_p, \mu_{ap})$ using NLS, controlling for the market-year effects $\{\gamma_{n_{ijt}}\}$ using the Frisch-Waugh-Lovell theorem.³³

³¹Note that (14) is compatible with differences in value added per worker across affiliates. These may be due to either differences in wages P_{ijt}^l or capital usage K_{ijt} . Wage variation across affiliates may reflect differences either in labor supply or labor market frictions, as affiliates may operate in different labor markets even within the same country-sector pair n . And, because we do not restrict the correlation between affiliate performance ψ_{ijt} and wages P_{ijt}^l , our model is consistent with differential labor frictions across affiliates depending on performance ψ_{ijt} (e.g. due to differences in screening or hiring mechanisms). Differences in K_{ijt} may reflect variation in any affiliate-specific state variable in \mathbf{S}_{it} , including, for example, performance ψ_{ijt} , the capital price P_{ijt}^k , or capital adjustment costs χ_{ijt}^k .

³²Theoretically, an alternative way of handling measurement error in value added would be to use a control function for unobserved performance ψ_{ijt} that exploits variation in either capital (Olley and Pakes 1996) or materials (Levinsohn and Petrin 2003), estimating the production function parameters either in the original form suggested in these papers or using the alternative approach in Akerberg, Caves, and Frazer (2015). In practice, a control function based on capital, a dynamic input, would require a nonparametric projection of va_{ijt} on all variables in the firm- i state vector; this is infeasible in our setting due to the large dimensionality of \mathbf{S}_{it} in (10), which precludes following Olley and Pakes (1996). We cannot use materials as a control variable, as we do not observe affiliate materials expenditures. This precludes following Levinsohn and Petrin (2003) and Akerberg, Caves, and Frazer (2015).

³³See Frisch and Waugh (1933), Lovell (1963), Giles (1984). It is important to have many affiliates operating in each market-year so that, after controlling for $\{\gamma_{n_{ijt}}\}$, the variation that remains is sufficient to identify the unknown parameters in (16). If variation within a market-year among the observed covariates in (16) is low, large standard errors for our estimates of $(\beta_k, \beta_{kk}, \rho, \mu_a, \mu_p, \mu_{ap})$ would result.

This second step identifies the impact of affiliate and parent R&D investment on affiliate performance ψ_{ijt} by exploiting variation in r_{ijt-1} and r_{i0t-1} conditional on lagged affiliate performance ψ_{ijt-1} and market-year unobserved effects $\gamma_{n_{ijt}}$. Controlling for ψ_{ijt-1} ensures that our estimates are not affected by reverse causality. Specifically, our estimates do not reflect the correlation between lagged R&D r_{ijt-1} and current performance ψ_{ijt} that would arise if performance is persistent and r_{ijt-1} is determined by ψ_{ijt-1} . In addition, market-year fixed effects $\gamma_{n_{ijt}}$ account for country characteristics that may affect both affiliate performance growth and affiliate incentives to perform R&D. For example, countries with an abundance of skilled labor may be attractive locations for affiliate R&D investment, and affiliates in such countries may also have higher levels of performance growth. Importantly, the fixed effects we include account for the possibility that such country characteristics may have differential effects on affiliates depending on their industry. All remaining variation that is unobserved by the econometrician, η_{ijt} , is assumed to be mean independent of the information known to the firm at period $t - 1$, \mathbf{S}_{it-1} , and is therefore uncorrelated with R&D investments at period $t - 1$.³⁴

Intuitively, our approach thus estimates the impact of R&D on future performance by comparing two affiliates in the same host country that share equal performance levels at $t - 1$ and that differ in parent and affiliate R&D investments, r_{i0t-1} and r_{ijt-1} . Conditional on affiliate performance and a market-year, note that optimal R&D investments of the affiliate and its parent may differ due, for example, to differences in their realized R&D cost shocks χ_{ijt-1}^r and χ_{i0t-1}^r (see section 2.6). Identification does not require fully specifying the impact of these R&D cost shocks on optimal R&D decisions, as we rely on observed R&D choices in the data for identification. Hence, our estimator is robust to alternative models of the R&D decision, as long as they are consistent with the assumption that R&D decisions at $t - 1$ are mean independent of performance shocks at t . In particular, the model in section 2.6 assumes that multinational firm i determines innovation investments optimally for each of its sites \mathcal{J}_{it} , but our estimation approach remains valid if each affiliate instead determines independently its own innovation investment, provided these investments are orthogonal to subsequent performance shocks.

This two-step estimation procedure yields consistent and asymptotically normal estimates of the parameters $(\beta_l, \beta_{ll}, \beta_{lk}, \beta_k, \beta_{kk}, \rho, \mu_a, \mu_p, \mu_{ap})$ under the assumptions in section 2.³⁵

³⁴Our estimation approach implicitly assumes that affiliates owned by parents that share a common sector have identical production parameters, even if the affiliates themselves operate in distinct industries. If these parameters were instead heterogeneous and correlated with the R&D investment of the affiliate or its parent, our estimates of (μ_a, μ_p, μ_{ap}) would be biased. To explore the importance of this concern, in section 6.4 we restrict the estimation sample to include only affiliates producing in a single industry (that of the parent), and obtain estimates quantitatively similar to our baseline results. An incipient literature (Kasahara, Schrimpf, and Suzuki 2015, Balat, Brambilla and Sasaki 2016) explores the extent to which it is feasible to permit unobserved heterogeneity in production parameters.

³⁵Consistency and asymptotic normality are guaranteed provided (14), (15), (16), and (17) hold. Equation (14) is derived from the demand function in (1), the production function in (2) and the assumptions that affiliates are monopolistically competitive and both materials and labor are static inputs. Given (14), (15) is implied by (7). Equation (16) is implied by (13), which is itself a function of the demand and production functions, the assumption that affiliates are monopolistically competitive, and the stochastic process for firm performance in (8) and (9). Finally, (17) is guaranteed by the mean independence between η_{ijt} and the state vector of firm i at period $t - 1$, \mathbf{S}_{it-1} , implied by (8) and our assumption in section 2.6 that decisions regarding both period- t capital k_{ijt} and whether affiliate j is active in period t , $j \in \mathcal{J}_{it}$, are taken in period $t - 1$.

4.2 Instantaneous Entry and Exit

The model presented in section 2 assumes affiliate entry and exit decisions taken at t are implemented after period- t production occurs. We consider here the possibility that entry and exit decisions taken at t are instead instantaneous, taking place prior to period- t production decisions.

Condition (17) requires that the expectation of the performance shock η_{ijt} is zero across affiliates active at both $t - 1$ and t . For affiliate j to be active at both $t - 1$ and t , it must be the case that a) entry by j occurs at or before $t - 1$, and b) exit by j occurs after t . Instantaneous entry is thus not a concern as (8) implies entry events that occur prior to period t are independent of η_{ijt} .

However, instantaneous exit decisions may introduce a form of sample selection bias that affects the estimator in section 4.1. The restriction $j \in \mathcal{J}_{it}$ in (17) is satisfied when firm i optimally chooses to maintain control over affiliate j at t . In a model with instantaneous exit, this period- t decision for affiliate j depends on current performance ψ_{ijt} , itself a function of lagged innovation r_{i0t-1} and r_{ijt-1} , lagged firm performance ψ_{ijt-1} , and the performance shock η_{ijt} . Even if firm i determines r_{ijt-1} and r_{i0t-1} before observing η_{ijt} , the fact that the exit decision for j occurs after all three of these variables are observed by the firm implies η_{ijt} is correlated with r_{ijt-1} , r_{i0t-1} and ψ_{ijt-1} conditional on survival.³⁶ Specifically, this correlation between η_{ijt} and both r_{ijt-1} and r_{i0t-1} would be negative, placing a downward bias on NLS estimates of μ_a , μ_p , and μ_{ap} (see Appendix A.4 for details).^{37,38}

5 Main Results

Estimates of our baseline model appear in Table 3.³⁹ As a preliminary step, we first evaluate the influence of innovation on affiliate performance in a specification that considers an affiliate as responsive only to its own R&D investment (column 1). We find evidence that the performance of an affiliate increases systematically in its own R&D investment and is also persistent, results that qualitatively match recent estimates (Aw, Roberts and Xu 2011, Doraszelski and Jaumandreu 2013). In column 2, we add R&D performed by the U.S. parent of the affiliate, and find strong support for its importance as a determinant of affiliate performance. In column 3, by contrast, we find that the innovation investment of other foreign affiliates within the same firm has no significant impact on affiliate performance, suggesting the strong centrality of parent innovation within the firm network.⁴⁰ In column 4, which includes the interaction between parent and affiliate R&D,

³⁶By contrast, while exit decisions taken at t are also a function of ψ_{ijt} in the baseline model with delayed exit, this does not restrict the set of period- t observations used for estimation, as choices take effect in the subsequent period. Thus, endogenous exit does not generate sample selection bias for the model in section 2.

³⁷Theoretically, one may correct this bias following the procedure in Olley and Pakes (1996). In our context, implementing this would require nonparametrically projecting the exit indicator for each firm- i affiliate j at t on the full firm- i state matrix \mathbf{S}_{it} . This projection is infeasible due to the high dimensionality of \mathbf{S}_{it} .

³⁸A distinct form of sample selection bias would arise if the multinational firm were to exit. While exit of multinational firms is minimal in our data, it is also important to observe that idiosyncratic affiliate-specific shocks are unlikely to determine the exit decisions of the overall multinational firm; as a result, this consideration is unlikely to be an important source of selection bias in our estimates.

³⁹Estimation details appear in Appendix A.9.

⁴⁰When interpreting these results, it is important to observe that our sample includes only manufacturing affiliates, and thus excludes foreign sites devoted purely to innovation. The innovation output of such sites could thus have an

we find evidence that relatively innovative affiliates also receive correspondingly higher gains from parent innovation.

The estimates in columns 2 and 4 reveal that the intrafirm impact of U.S. parent R&D investment on affiliate performance is economically significant. All else equal, the estimates in column 2 indicate that the elasticity of affiliate performance with respect to parent R&D investment μ_p is approximately four times larger than the elasticity μ_a with respect to affiliate R&D spending. The large estimated magnitude $\hat{\mu}_p$ is of interest from an innovation policy perspective, as it suggests that the welfare effects of innovation stimulus incentives such as R&D tax credits may extend to firm sites located abroad. Our estimates suggest that such policies, where effective at increasing R&D investment by multinational parents (Rao 2016), increase affiliate performance abroad and may thereby contribute to a spatial disconnect between the costs and gains of such policies. From the perspective of an affiliate host country, these estimates also quantify potential gains due to local manufacturing and innovation activity by affiliates of foreign multinational firms.

Notice that the estimated differences between μ_a and μ_p in column 2 are nevertheless compatible with an optimizing firm that equalizes the marginal net return to its parent and affiliate R&D investments. The assumptions in section 2 imply that while the marginal gross return to R&D investment by affiliate j increases in μ_a , it decreases in the R&D volume R_{ijt-1} ; similarly, the marginal gross return to R&D by the firm- i parent increases in μ_p and decreases in R_{i0t} . Thus, given that $\hat{\mu}_p > \hat{\mu}_a$ in column 2, an optimizing firm facing similar marginal costs of R&D investment across its sites would perform more U.S. parent R&D than affiliate R&D. Parents indeed perform much higher R&D volumes than foreign affiliates in the data (section 3).⁴¹ In addition, comparing the μ_a estimates in columns 1 and 2 also suggests an important consequence of failing to account for parent innovation, as in column 1. Specifically, this omission leads to an overstatement—of approximately 50 percent—in the performance impact of affiliates’ own R&D, reflecting a positive correlation across firms between parent and affiliate R&D investment.

The complementarity between parent and affiliate innovation that we find in column 4 implies innovating affiliates benefit more from U.S. parent R&D than non-innovating affiliates: the impact of parent R&D on non-innovating affiliates is just 60 percent as large as its average impact in column 2. This result supports the view put forth in Cohen and Levinthal (1989), which proposes that, beyond its potential for generating new production techniques, R&D enhances a firm’s ability to assimilate and exploit existing information—information developed by the parent firm, in this case. Our estimates shed new light on this theoretical possibility, and indicate that an affiliate’s gain from parent technology is amplified by its own investment in R&D.

These results also contribute to our understanding of affiliate and parent roles within a multinational firm. Through the positive estimates of μ_a in columns 1–3 and the positive interaction μ_{ap} in column 4, we find that affiliates’ contribution to proprietary technology development has meaningful consequences for performance outcomes; this adds nuance to the standard view of the

impact on firm performance not reflected by our estimates.

⁴¹Notice, however, that our model does not impose equal innovation costs across sites; the R&D cost shifters χ_{ijt}^r included in (10) thus also contribute to the distribution of optimal R&D spending across firm sites through differences in sites’ fixed and variable R&D costs. See footnote 24 for additional details.

multinational firm (Helpman 1984, Markusen 1984), which considers affiliates as pure recipients, rather than producers, of knowledge. Our results also provide a rationale for heterogeneous, yet correlated, affiliate productivities within the multinational firm, as specified in Arkolakis et al (2014). This standard view also emphasizes a range of firm-specific assets including not only R&D investment, but also marketing and management, that are produced centrally and shared with affiliates within the boundaries of the firm. Among these, our estimates reveal and quantify the importance of parent R&D investment specifically. Given the prevalence of government innovation policies aimed at stimulating private R&D investment at the country level, understanding how the gains from this activity are distributed *across countries* is of interest from a policy perspective. Our estimates indicate that parent innovation is central, significantly increasing foreign affiliate performance within the firm. In what follows, we therefore emphasize specifications like columns 2 and 4 that include both own-affiliate and parent R&D spending.

The lower rows of Table 3 present the mean and standard deviation of the distribution, across affiliates, of value added elasticities with respect to labor and capital. Our model predicts that these elasticities β are strictly smaller than the corresponding output elasticities α , and are thus compatible with approximately constant returns to scale.⁴²

6 Robustness and Alternative Specifications

We consider a number of alternative specifications to better establish the stability of our main results, and to thereby shed light on the importance of the assumptions in our baseline model.

6.1 Intrafirm Technology Licensing

One interpretation of observed technology royalties and license fees flowing within the multinational firm is that they are an exact proxy for otherwise unobserved technology transfer (Hines 1995, Branstetter, Fisman, and Foley 2006). Under this interpretation, parent R&D should have no effect on affiliate performance once we control for these observed technology royalties and license fees. We can therefore use our estimating equation, extended to include royalty payments and technology license fees, to evaluate the extent to which these capture technology transfer.

Corresponding estimates appear in Table 4, and indicate that the impact of parent R&D on affiliate performance is positively correlated with royalties paid by affiliates to parents, consistent with the view that these royalty payments and license fees capture meaningful intrafirm technology links. Specifically, estimated coefficients on royalty payments in columns 1 and 2 are positive and

⁴² As $\sigma > 1$ and $0 < \alpha_m < 1$, the parameters of the value added function β in (5) are always smaller in absolute value than the corresponding production function parameters α . Notice that for an elasticity of substitution σ around 5 (e.g. Head and Mayer 2014) and a materials expenditure share in sales α_m around 0.65 (Doraszelski and Jaumandreu 2013), the scale parameter ι defined in section 2.4 is approximately 1.6. From (2), the implied average output elasticities are thus: a) $0.468\iota(1 - \alpha_m) = 0.468 \times 1.6 \times (1 - 0.65) = 0.26$ with respect to labor; and b) $0.253 \times 1.6 \times (1 - 0.65) = 0.14$ with respect to capital. Summing these with α_m yields a coefficient near 1, implying that our estimates are consistent with the average affiliate having a constant returns to scale production function. Notice however that the translog production function implies that each affiliate may have a different labor and capital elasticity in equilibrium. Our estimates indicate that the coefficient of variation for the labor elasticity is approximately 0.12, and for the capital elasticity is approximately 0.2.

statistically significant; in addition, coefficients on U.S. parent R&D investment fall relative to the baseline estimates in Table 3, suggesting that these payments indeed reflect performance gains resulting from the acquisition of technology developed by the U.S. parent. However, columns 3 and 4 of Table 4 indicate that R&D performed by the U.S. parent also impacts affiliates that *do not* report paying royalties or license fees to their parent firm, particularly for innovating affiliates, suggesting an effect of parent innovation not fully captured by royalty payments.

6.2 Alternative Performance Process

We assess the sensitivity of our estimates to several aspects of the affiliate performance process. First, we assume in (9) that the direct impact of innovation investment on affiliate performance occurs one year after the investment date. However, certain technology development projects may require additional time before the impact on affiliate performance is realized. To explore this possibility, we estimate variants of (9) in which r_{ijt-1} is replaced by either r_{ijt-2} or r_{ijt-3} , similar to Aw, Roberts, and Xu (2008). The resulting estimates appear in columns 1 through 4 of Table 5 and are quantitatively similar to those in Table 3. The main estimates are thus not sensitive to this adjustment in the assumed timing of innovation impact.⁴³

Second, section 3 indicates that, while almost all U.S. parents perform R&D, only one quarter of foreign affiliates do so. We therefore estimate an alternative model that relaxes the assumption in (9) that all affiliates, even those choosing the corner solution of zero innovation spending, are affected equally by the market-year factors accounted for by the performance process effects $\{\mu_{n_{ijt}}\}$. Specifically, we extend (9) to include separate market-year effects that correspond to innovating affiliates: these additional controls interact market-year fixed effects with d_{ijt-1} , an affiliate-specific indicator variable that is equal to one if firm i 's affiliate j performs R&D during $t-1$. These effects allow for flexible differences in average performance growth between innovating and non-innovating affiliates located within the same market and year.⁴⁴ The resulting estimates appear in columns 5 to 8 of Table 5. Relative to the baseline results in section 5, those in Table 5 are nearly identical regarding the affiliate performance elasticity with respect to parent innovation. However, the elasticity with respect to an affiliate's own innovation μ_a is larger in Table 5 than our baseline estimate. Moreover, a consequence of this increase in μ_a is that it is now similar in magnitude to μ_p , the elasticity of affiliate performance with respect to parent R&D investment (column 6).

Third, the performance process in (9) imposes a strong Markov assumption: conditional on current affiliate performance and R&D investments, expected future performance is independent of its prior path. If this condition is not satisfied in the data, the error term in (16) will include affiliate performance terms in periods prior to $t-1$; these may be correlated with the measures of R&D investment included as covariates in (16). Specifically, serial correlation in affiliate R&D implies omitting relevant lags of affiliate performance would cause upward bias in the estimates of

⁴³In part, this reflects persistence in affiliate-level innovation spending. Such persistence makes identifying the precise lag with which innovation investments translate into productivity increases impossible. Consequently, in section 7 we emphasize the long-run implications of our estimates, considering a steady state in which site-level innovation investments are assumed to be constant over time.

⁴⁴To obtain estimates that are independent of the scale with which R_{i0t} and R_{ijt} are measured in the data, we could further extend the model to include the interaction between d_{ijt-1} and parent R&D r_{i0t-1} .

μ_a , μ_p and μ_{ap} reported in section 5. We evaluate the sensitivity of our baseline results to this possibility by expanding (9) to include additional performance lags ψ_{ijt-2} and ψ_{ijt-3} ; resulting estimates appear in columns 1 through 4 of Table 6. While the coefficient on affiliate R&D weakens, those on parent R&D and its interaction with affiliate R&D, μ_p and μ_{ap} , remain essentially unchanged. Including these additional lags helps control for the possible presence of unobserved, serially correlated determinants of performance (e.g. management) when these are endogenously determined as a function of further performance lags. Thus, the observed invariance of our μ_p and μ_{ap} estimates to including ψ_{ijt-2} and ψ_{ijt-3} indicates that any parent innovation responsiveness to unobserved idiosyncratic variables is likely to be small. The traditional endogeneity bias affecting estimated productivity impacts of R&D in single-plant firm settings therefore seems unlikely to be driving our estimates of the affiliate-level impact of parent R&D.

Finally, a multinational firm i may experience shocks that simultaneously affect the performance of all firm- i affiliate sites. Columns 5 and 6 of Table 6 provide estimates that allow performance shocks η_{ijt} to be correlated across affiliates within the same multinational firm and year. The results are qualitatively similar to those in columns 2 and 4 of Table 3.

6.3 Labor Measurement

As discussed in section 3, our baseline results in Table 3 measure labor inputs L_{ijt} as the total number of workers employed by affiliate j . However, because an innovating affiliate may devote some workers to innovation, a more precise measure of L_{ijt} would be the total number of production workers. Information on the division of labor between innovation and production is available in benchmark years 1989, 1994, 1999, and 2004. We use this information to compute an affiliate-specific share of workers employed in innovation and apply it across all sample years to construct a measure of production workers.⁴⁵ Estimates using this alternative measure appear in columns 7 and 8 of Table 6, and are similar to the baseline results in Table 3.⁴⁶

6.4 Heterogeneous Innovation Impact

While the results in section 5 evaluate a model with identical parameters governing affiliate production, demand, and performance evolution across foreign affiliates of U.S. parent firms operating in the computer industry, the impact of parent R&D on affiliate performance may differ systematically depending on the industrial proximity between an affiliate and its parent. Similarly, the affiliate production function may depend on the output industry of the affiliate. Columns 1 and 2 in Table 7 thus restrict the sample to include only those affiliates operating within the computer industry (SIC 357), the same sector as the U.S. parent, to assess the importance of this consideration. The resulting estimates suggest that restricting the composition of affiliates in this way significantly increases the complementarity between parent and affiliate innovation, but does not affect the impact of parent R&D on the performance of affiliates not investing in R&D.

The computer industry has also evolved during our sample period. In particular, the prolif-

⁴⁵The data unfortunately do not permit an analogous approach for capital inputs. See footnote 30 for discussion.

⁴⁶Appendix A.3 contains additional details regarding the construction of our measure of production workers.

eration of personal computers and other devices could imply that the degree of substitutability between computer products available in 2008 could be higher than that in 1989. In Appendix A.5, we show that such heterogeneity in the demand elasticity σ would imply that the performance effects of parent and affiliate R&D are heterogeneous. In columns 3 through 6 of Table 7, we therefore reestimate the model parameters allowing the demand elasticity parameter σ to vary by year to account for the possibility of such evolution. Following Appendix A.5, we obtain estimates that are indeed consistent with σ increasing over time. To determine the importance of this for our results, we compute the distribution of the affiliate performance elasticity with respect to R&D spending. Comparing the estimates in columns 1 through 4 in Table 3 with the elasticity distributions implied by the estimates in Table 7, we find that those computed under the homogeneous- σ assumption are near the averages elasticities in which σ may vary by year.⁴⁷

6.5 Other Industries

Table 8 assesses the relevance of the results above for multinationals in other industries. Columns 1 through 8 in Panel A show that the essential patterns observed in Table 3 are upheld within multinational firms in pharmaceutical drugs (SIC 283) and in the motor vehicles industry (SIC 371). One difference is that, in the motor vehicles sector, the coefficient on the interaction between parent and affiliate R&D is not statistically distinguishable from zero. To determine whether our results are sensitive to the industry definition, we reestimate the model parameters within each of the three associated broader sectors: industrial machinery (SIC 35), chemicals (SIC 28), and transportation equipment (SIC 37).⁴⁸ Table 8 Panel B shows that the main qualitative results are upheld, and are thus not sensitive to the degree of industry aggregation in the sample.

7 Quantitative Implications

In this section, we investigate implications of the estimates described in section 5. We quantify the contribution of headquarters innovation to long-run affiliate performance, to the overall value added earned by a multinational firm, and to the aggregate value added earned in countries hosting affiliates. Throughout this section, we rely on the estimates reported in column 4 of Table 3.⁴⁹

7.1 Innovation and Affiliate Performance

The welfare gains from trade with multinational production hinge on multinationals' productivity distribution across the set of production locations (Arkolakis et al 2014). Our estimates imply that this productivity distribution is influenced by parent and affiliate innovation, and the respective impacts of these investments on each production site within the multinational firm.

⁴⁷To characterize the distribution across years for each elasticity, one should multiply the corresponding estimate in columns 3–6 of Table 7 by the quantiles of the distribution of ι_{2t} in the last column. For example, according to column 3, the median elasticity of affiliate performance with respect to lagged affiliate R&D is $0.0021 \times 2.84 = 0.0059$, the minimum is $0.0021 \times 2.07 = 0.0043$, and the maximum is $0.0021 \times 3.72 = 0.0078$. See Appendix A.5 for details.

⁴⁸In the interest of space, we omit results for transportation equipment (SIC 37); these are available upon request.

⁴⁹Section 8 below evaluates the same set of implications using estimates that permit intrafirm parent-affiliate trade.

To quantify the long-run impact of parent and affiliate innovation on affiliate performance, we consider their effects on $\psi_{ij} \equiv \mathbb{E}[\lim_{s \rightarrow \infty} \psi_{ijs} | r_{i0t}, r_{ijt}]$, the expected long-run performance of affiliate j within firm i ; for simplicity, we consider the benchmark case in which we fix firm i 's parent and affiliate- j innovation levels r_{i0t} and r_{ijt} at their respective period- t values for a specified base year t . The long-run performance impact of eliminating the contribution of parent innovation can then be evaluated by considering how this expectation responds to setting $r_{i0t} = 0$. Holding affiliate- j innovation fixed, this expectation becomes $\psi_{ij,r_0} \equiv \mathbb{E}[\lim_{s \rightarrow \infty} \psi_{ijs} | 0, r_{ijt}]$, and the impact of removing the contribution of parent R&D is therefore $\Delta_{ij,r_0} \equiv \psi_{ij,r_0} - \psi_{ij}$.⁵⁰ The long-run performance impact of eliminating affiliate R&D may be expressed as the analogous Δ_{ij,r_j} .

Figure 1 plots the distributions of Δ_{ij,r_0} and Δ_{ij,r_j} across affiliates using data for the base year $t = 2004$.⁵¹ At each percentile, it is apparent that $\Delta_{ij,r_0} < \Delta_{ij,r_j}$, indicating that long-run affiliate performance is lowest in the absence of parent innovation.

These distributions reveal the economic significance of parent innovation for foreign affiliates, reflecting not only the elasticities presented in Table 3, but also the underlying distributions of innovation investment across parents and affiliates in the data. For the median affiliate, Figure 1 indicates that eliminating the impact of parent R&D would, all else equal, imply a substantial reduction in its long-run performance (62.3 percent); the reduction due to eliminating the affiliate's own innovation is smaller by a factor of three (17.7 percent). While the relative impacts of removing these two sources of performance gain differ across affiliates, parent innovation is always more important. Specifically, Figure 1 reveals a decline in long-run affiliate performance of 80 to 90 percent for the quintile of affiliates most impacted by losing access to parent R&D. The analogous decline due to eliminating affiliates' own innovation is only 40 to 60 percent.⁵²

The evident importance of parent innovation for the long-run performance of affiliates within the same multinational firm sheds light on the quantitative importance of intangible input transfer across firm sites. In particular, our results suggest that ownership theories emphasizing the intrafirm transfer of technological inputs are relevant for explaining the existence of multinational firms, and also provide a rationale for evidence indicating affiliates acquired by foreign multinationals are faster growing than unaffiliated firms.⁵³

7.2 The Headquarters Performance Advantage

Recent work including Tintelnot (2016) and Head and Mayer (2016) finds that parents are systematically more productive than foreign affiliates within the multinational firm. While this performance

⁵⁰Accounting for the impact of a change in parent innovation on optimal affiliate R&D would require fully specifying the dynamic problem of firm i in section 2.6 and solving the optimization in (11). As described in footnote 24, this approach would pose a severe computational challenge, and would also require imposing a set of assumptions on firm behavior that our analysis does not otherwise need.

⁵¹See Appendix A.10 for details. We replicate these calculations with the alternative base year $t = 1994$ and find similar results. Both 1994 and 2004 are among the BEA surveys that are the most comprehensive in coverage.

⁵²If we instead base these computations on the estimates in columns 5 through 8 of Table 5, the conclusions resulting from Figure 1 remain essentially unchanged. This is because, although the importance of affiliate innovation increases relative to that of parent R&D, Figure 1 hinges on the distribution of parent and affiliate R&D spending in the data.

⁵³For example, Arrow (1975), Teece (1982), Atalay, Hortaçsu, and Syverson (2014); Arnold and Javorcik (2009), Guadalupe, Kuzmina, and Thomas (2012), and Javorcik and Poelhekke (2016).

advantage likely reflects a range of determinants, the framework in section 2 may be combined with our estimates to quantify the specific contribution of innovation investment to the measured parent-affiliate performance gap observed within firms in the data.

We measure this intrafirm performance gap by comparing the expected long-run performance of each parent firm in the sample with that of its foreign affiliates. This requires first estimating the determinants of parent performance evolution; we proceed by assuming parents face demand and production functions similar to (1) through (4), but governed by distinct parameter values σ and α . We also assume parent performance evolves according to a version of (8) with

$$\mathbb{E}_{t-1}[\psi_{i0t}] = \rho_0\psi_{i0t-1} + \mu_0r_{i0t-1} + \mu_{a0}r_{i-0t-1} + \mu_{n_{i0}t}, \quad (18)$$

instead of (9), where ρ_0 is the persistence of parent performance, μ_0 is the impact of parent R&D, and μ_{a0} is the impact of r_{i-0t-1} , the log sum of R&D investment across firm- i affiliate sites. NLS estimates of ρ_0 , μ_0 , and μ_{a0} appear in Table A.3. These suggest parent innovation is an important determinant of its own performance, while R&D performed by its foreign affiliates is not.⁵⁴

The relative long-run performance difference $\psi_{ij} - \psi_{i0}$ between the parent of firm i and its affiliate j depends, in the model, on the respective sequences of market-year specific performance shocks affecting these two sites, $\{\mu_{n_{i0}s}, s \in [t, \infty)\}$ and $\{\mu_{n_{ij}s}, s \in [t, \infty)\}$. Given the data described in section 3, these unobserved effects cannot be separately identified from the market-year unobserved terms $\{\kappa_{n_{i0}s}, s \in [t, \infty)\}$ and $\{\kappa_{n_{ij}s}, s \in [t, \infty)\}$ in the parent and affiliate value added functions. We therefore focus here on the parent-affiliate performance gap due exclusively to the differential impact of parent and affiliate innovation. Appendix A.11 shows how this may be evaluated using affiliate-level estimates from section 5, parent-level estimates from Table A.3, and data for a base year t . Figure 2 shows the resulting distribution of the difference $\psi_{ij} - \psi_{i0}$ for the base year $t = 2004$. The average affiliate reaches a long-run performance level that is only 51 percent as high as its U.S. parent, and within the sample, all affiliates exhibit lower long-run performance levels due to innovation than their respective parents.⁵⁵ Our estimates thus indicate that multinationals' endogenous decisions regarding site-specific innovation are able to explain a substantial performance advantage for parents relative to affiliates within multinational firms.

7.3 Innovation Policy

The U.S. Congress distributes over \$30 billion in innovation subsidies to private U.S. firms annually, to offset costs of R&D investment performed in the United States (National Science Board

⁵⁴Table A.3 also explores the possibility that the impact of affiliate R&D on parent performance depends on the affiliate industry. Considering separately the impact of R&D by affiliates producing in the same 3-digit industry as the parent yields an estimate that remains statistically insignificant.

⁵⁵To build intuition, consider a simpler setting with $\mu_a = \mu_{ap} = 0$. Comparing the estimates in Tables 3 and A.3, note that the short-run elasticity μ_p of affiliate performance with respect to parent R&D is larger than the corresponding elasticity μ_0 of parent performance. However, the higher persistence of parent performance ($\rho_0 > \rho$) implies the opposite is true in the long-run: the elasticity of long-run parent performance with respect to parent R&D is $\hat{\mu}_0/(1 - \hat{\rho}_0) = 0.0106/(1 - 0.918) = 1.29$, while that of affiliates' long-run performance is only $\hat{\mu}_p/(1 - \hat{\rho}) = 0.0164/(1 - 0.772) = 0.72$. Thus, if parent R&D were the only determinant of performance, parents would be 75 percent more productive than their affiliates in the long run. The difference between this number and the distribution in Figure 2 reveals the influence of affiliate innovation on the performance gap.

2014). Because large corporations including U.S.-based multinationals claim the majority of these subsidies, their impact on aggregate R&D hinges on the responsiveness of U.S. parent innovation spending to changes in innovation costs. The U.S. parent innovation response to a subsidy depends on the multinational firm-wide return to parent R&D investment. We quantify the share of this firm-wide R&D return that accrues to foreign affiliates, given our estimates and underlying data, and find that it is substantial; this suggests that failing to account for the impact of parent innovation on foreign affiliates not only understates the firm-wide return to parent innovation, but also distorts the predicted efficacy of U.S. innovation policy.

Provided that parents of multinational firms determine innovation investment optimally, and that firm- i parent R&D R_{i0t} is positive, it must satisfy

$$\frac{\partial V(\mathbf{S}_{it})}{\partial R_{i0t}} = \underbrace{\frac{\partial \sum_{s>t} \mathbb{E}_t \left[\sum_{j \in \mathcal{J}_{is}} \delta^{s-t} V A_{ijs}^* \right]}{\partial R_{i0t}}}_{\equiv GR_{i0t}} - \underbrace{\frac{\partial C_r(R_{i0t}, \chi_{i0t}^r)}{\partial R_{i0t}}}_{\equiv MC_{i0t}} = 0, \quad (19)$$

where GR_{i0t} is the gross firm- i return to parent R&D and MC_{i0t} is its marginal cost; as defined in section 2, $V(\cdot)$ is the firm- i value function, \mathbf{S}_{it} is the firm- i state vector, VA_{ijt}^* is the true value added of firm i 's affiliate j , and $C_r(R_{i0t}, \chi_{i0t}^r)$ is the parent R&D cost function, itself a function of R_{i0t} and the firm- i parent R&D cost shock χ_{i0t}^r .⁵⁶ Optimal firm- i parent R&D spending thus depends on the shape of the gross return function GR_{i0t} ; moreover, firms for which the function GR_{i0t} is larger at any given level of parent R&D will optimally choose higher levels of parent innovation investment, all else equal. The model in section 2 and the estimates in section 5 imply that, at any given level of parent R&D, GR_{i0t} is increasing in: a) the number and total value added of firm- i affiliates at t , b) the R&D expenditure of firm- i foreign affiliates at t , c) expected growth at t in the variables in a) and b). We quantify the combined contributions of a) and b) to the gross firm- i return to parent R&D, assuming for simplicity that firm i expects the number of affiliates, affiliate value added, and R&D spending to remain stable at period- t levels.⁵⁷ In this case,

$$GR_{i0t} = \frac{\partial \mathbb{E}_t[\psi_{i0t+1}]/\partial r_{i0t}}{1 - \rho_0} \frac{VA_{i0t}^*}{R_{i0t}} + \sum_{\substack{j \in \mathcal{J}_{it} \\ j \neq 0}} \frac{\partial \mathbb{E}_t[\psi_{ij t+1}]/\partial r_{i0t}}{1 - \rho} \frac{VA_{ijt}^*}{R_{i0t}}, \quad (20)$$

where the first term captures the return to firm- i parent innovation that is realized by the parent, the second term captures the return realized by its foreign affiliates, and where according to (18) and (9), $\partial \mathbb{E}_t[\psi_{i0t+1}]/\partial r_{i0t} = \mu_0$ and $\partial \mathbb{E}_t[\psi_{ij t+1}]/\partial r_{i0t} = \mu_p + \mu_{ap} r_{ijt}$.

The distribution of GR_{i0t} across firms appears in Figure 3 for the base year $t = 2004$. The average firm-wide gross return is 2.1 dollars and the median is 0.50 dollars.⁵⁸ The share of the

⁵⁶Parent R&D is almost always positive in our estimation sample (see section 3). As in Doraszelski and Jaumandreu (2013), the function $C_r(\cdot)$ accounts for the possibility that R_{i0t} may not capture all costs associated with R&D activity. This is consistent with the data, as our measure of R&D spending excludes capital innovation costs.

⁵⁷Detailed derivations appear in Appendix A.12.

⁵⁸Notice that we have not imposed the equality in (19) when computing the estimates in section 5, yet the estimates in section 5 indicate that the gross long-run return is in the vicinity of one for most firms. This implies that the effective marginal cost of investing one additional dollar in R&D is also near one for most firms.

gross return to firm- i parent innovation that is realized among its foreign affiliates is

$$\Lambda_{it} \equiv 1 - \left(\frac{\mu_0}{1 - \rho_0} \frac{VA_{i0t}^*}{R_{i0t}} \right) / GR_{i0t},$$

and the distribution of Λ_{it} appears in Figure 4 for the base year $t = 2004$. The average value of Λ_{it} across firms is 27 percent, implying that the gross private return to parent R&D investment exceeds the parent-level return by approximately 37 percent for the average firm, and that estimates of the innovation return based only on firms' domestic data are therefore understated substantially.

We find that the affiliate share in firm value added varies less across multinational firms than the affiliate share in value added gains from parent R&D Λ_{it} ; the latter also exceeds the former in the majority of firms.⁵⁹ The distinction between these distributions reflects, in part, the influence of affiliate R&D investment, which amplifies value added gains from parent R&D.⁶⁰

The importance of affiliates in the firm-wide return to parent innovation implies, through (19), that multinationals with more extensive affiliate operations optimally choose higher levels of parent R&D investment, as do firms with more innovative affiliates. This has implications for the responsiveness of parent innovation to U.S. R&D policy incentives. For example, U.S. R&D tax credits that subsidize U.S. innovation by private firms, captured by the model in section 2 through the parent R&D cost shifter χ_{i0t}^r , impact optimal firm- i parent R&D to an extent that hinges on the gross return function (20), and thus on the foreign affiliate presence of the firm.⁶¹ Specifically, the impact of a policy change in χ_{i0t}^r on optimal parent innovation by firm i is increasing in $\partial GR_{i0t} / \partial R_{i0t}$, the derivative of the gross return with respect to parent R&D. Because this derivative itself reflects the influence of parent innovation on affiliate sites abroad, through μ_p and μ_{ap} , responsiveness to U.S. innovation policy depends on firms' affiliate presence in markets abroad.

The multi-country nature of the gross return function also reinforces the spatial disconnect between the location of innovation financed by subsidies that affect χ_{i0t}^r and the location of innovation returns that enter GR_{i0t} . While this disconnect is internalized by firm i in our model, it is likely not internalized by U.S. policy makers determining the generosity of innovation subsidies, and could thus result in the underprovision of innovation subsidies.⁶²

7.4 U.S. Parent Innovation and GDP Growth Abroad

The affiliate-level estimates in section 5 indicate that parent innovation, through its positive impact on affiliate performance, raises the value added earned by foreign affiliates within the same

⁵⁹The affiliate share in total firm value added is $\Lambda_{it}^{va} \equiv \sum_{j \in \mathcal{J}_{it}, j \neq 0} VA_{ijt}^* / (VA_{i0t}^* + \sum_{j \in \mathcal{J}_{it}, j \neq 0} VA_{ijt}^*)$.

⁶⁰Intuitively, note that our model implies the distributions of Λ_{it} and Λ_{it}^{va} coincide if parent and affiliate performance are equally persistent ($\rho_0 = \rho$), and share the same elasticity with respect to parent R&D ($\mu_p = \mu_0$, $\mu_{ap} = 0$). In our estimates, these restrictions are not upheld, and it is thus not surprising that the distributions are distinct.

⁶¹Specifically, totally differentiating (19) with respect to χ_{i0t}^r and noting that $\partial GR_{i0t} / \partial \chi_{i0t}^r = 0$ yields

$$\frac{dR_{i0t}}{d\chi_{i0t}^r} = \frac{\partial MC_{i0t}}{\partial \chi_{i0t}^r} \left/ \left(\frac{\partial GR_{i0t}}{\partial R_{i0t}} - \frac{\partial MC_{i0t}}{\partial R_{i0t}} \right) \right. \quad (21)$$

Second order conditions imply that the denominator in (21) is negative. Any policy change negatively impacting the marginal cost of R&D will thus have a positive impact on optimal R&D spending. Therefore, the model in section 2 is consistent with U.S. R&D subsidies increasing R&D by parents, as found in Table A.2 (Appendix A.15).

⁶²See footnote 3.

multinational firm. An implication of this is that aggregate U.S. parent R&D investment raises total value added (or gross domestic product, GDP) in countries that host affiliates of U.S. firms.

To quantify the contribution of U.S. parent R&D to foreign countries' GDP, we consider the short-run impact of a policy change that prevents U.S. firm affiliates from receiving the intrafirm benefit of parent innovation.⁶³ Specifically, consider the period- $t + 1$ implications of such a policy implemented at t , holding fixed all other characteristics of parents, affiliates, and domestic firms.⁶⁴

We evaluate this impact using data for $t = 2004$. The results indicate that the aggregate value added earned by U.S. firm affiliates in $t + 1 = 2005$ falls by 18 percent for the average foreign country.⁶⁵ For the average country hosting U.S. firm affiliates, this decline is equal to 1.3 percent of its realized GDP in the computer industry. This impact is heterogeneous across countries and is increasing in the number and size of affiliates present and in affiliates' R&D investment; these correlate in the data with proximity to the United States, skill abundance, and the quality of intellectual property protection. For example, we find that GDP in the computer industry would decline in countries near the United States including Canada, by 7.47 percent, and Mexico, by 1.61 percent. The impact is also more pronounced for countries with a computer industry dominated by U.S. affiliates, including Finland (10.85 percent decline) and Spain (2.04 percent decline).

8 Intrafirm Trade

The model in section 2 considers horizontal multinational firms that do not engage in intrafirm input trade, consistent with Ramondo, Rappoport, and Ruhl (2016). However, even limited volumes of intrafirm trade could channel proprietary technology from parents to foreign affiliates (Keller and Yeaple, 2013; Irrazabal, Moxnes, and Oromolla, 2013). Imported inputs may also complement R&D investment (Bøler, Moxnes, and Ulltveit-Moe, 2015). To distinguish between tangible and intangible modes of technology transfer, and to assess possible complementarities between affiliates' intrafirm imports and R&D investment, we extend the specification in (9) to account for the volume of affiliate intrafirm imports from U.S. parents that is observed in the data.

For this extension of the model, we replace the affiliate performance process in (9) with

$$\begin{aligned} \mathbb{E}_{t-1}[\psi_{ijt}] = & \rho\psi_{ijt-1} + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \mu_{pm} r_{i0t-1} im_{ijt-1} \\ & + \mu_{apm} r_{ijt-1} r_{i0t-1} im_{ijt-1} + \mu_m im_{ijt-1} + \mu_{n_{ijt}}, \end{aligned} \quad (22)$$

where im_{ijt-1} is the log of imports received by firm i 's affiliate j from its U.S. parent at $t - 1$. Estimates appear in Table 9. Columns 1 and 2 show that parent and affiliate R&D remain significant determinants of affiliate performance growth after controlling for affiliate imports from the parent.

⁶³For a general-equilibrium evaluation of the theoretical mechanisms linking innovation, its responsiveness to innovation policy, and aggregate growth outcomes, see Atkeson and Burstein (2015).

⁶⁴We consider the impact before a) affiliates adjust labor, capital investment, and innovation, b) domestic firms in the affiliate host country adjust output, c) entry by either domestic firms or affiliates of non-U.S. multinationals occurs in the affiliate host country. Assessing the total impact of such a policy would also require accounting for possible spillover effects of resulting changes in U.S. affiliate performance on domestic plants (e.g. Javorcik 2004a), which could occur through links with domestic input producers (Rodríguez-Clare 1996). See also Keller (2002).

⁶⁵See Appendix A.13 for details.

The impact of intrafirm imports on affiliate performance is also positive and significant. In addition, the estimated impact of affiliate R&D is smaller in column 1 of Table 9 than in specifications that do not control for im_{ijt-1} (column 2, Table 3), consistent with the fact in the data that affiliates importing from their U.S. parent are also more likely to invest in R&D.

Columns 3 and 4 of Table 9 evaluate the hypothesis that parents share technology with affiliates by embedding it in physical goods. The positive and statistically significant estimate on the interaction between parent R&D and imports in column 3 supports this view: affiliates that import more from their parent benefit more on average from R&D performed by the parent. The lack of statistical significance in the estimated coefficient on parent R&D in column 3 further suggests technology is, to a certain extent, shared through tangible goods traded within the firm, though column 4 indicates this is true only for affiliates that do not perform R&D. Specifically, the estimated coefficient on the interaction between parent and affiliate R&D in column 4 indicates innovating affiliates benefit from parent innovation, even in the absence of imported inputs; intangible knowledge sharing may thus be particularly important for such affiliates.

8.1 Quantitative Implications

We replicate the quantitative investigation in section 7 using estimates that allow for intrafirm trade (Table 9, column 4). To assess the impact of parent innovation on long-run affiliate performance, we plot the distributions of Δ_{ij,r_0} , Δ_{ij,r_j} , and Δ_{ij,im_j} across affiliates in the base year $t = 2004$ in Figure 5. These again indicate that parent innovation is economically significant for firm affiliates: for the median affiliate, eliminating the impact of parent R&D would imply a 57.8 percent reduction in its long-run performance; analogous reductions due to eliminating the impact of parent imports and affiliate R&D are comparatively modest at 11.8 percent and 7.8 percent, respectively. Headquarters innovation is seven times as important as affiliate R&D and five times as important as imports from the parent in determining its long-run performance for the median affiliate.

In this extended setting, it remains the case that the impact of parent innovation determines a significant performance advantage for parents relative to affiliates within the multinational firm. In addition, a substantial share of the gross return GR_{i0t} to U.S. parent R&D is realized by foreign affiliates: Λ_{it} is 23 percent on average, comparable to the magnitude found in the baseline model. In summary, we find that adding imports to the analysis adds nuance, but does not substantially change the quantitative importance of U.S. parent R&D for affiliate performance.

9 Misreporting

A multinational firm may attempt misreporting affiliate profits to minimize its worldwide tax burden. To achieve this aim, a firm could misreport affiliate value added or affiliate R&D spending in response to prevailing corporate tax rates faced by its affiliates.⁶⁶ If prevalent in the data, it may bias the estimates in section 5 and affect the quantitative implications described in section 7.

Differences between actual and reported value added are accommodated by the model through

⁶⁶See Hines and Rice (1994), Hines (1997), and Bernard, Jensen, and Schott (2006).

the term ε_{ijt} in (6). Provided that these differences are uncorrelated with affiliate labor and capital input use, equation (15) is satisfied and the estimation procedure in section 4.1 will yield consistent estimates even in the presence of misreporting. In this section, we assess the robustness of our main results to patterns of value added misreporting that do not verify (7). Specifically, we re-estimate the model using two subsamples. First, we exclude affiliates located in tax havens (Gravelle 2015). All else equal, if misreporting of value added for tax purposes is present in our data, the set of affiliates excluded from this first subsample will tend to overreport value added. Second, we exclude affiliates importing a high share of inputs from their U.S. parent. All else equal, if value added misreporting is present, the set of affiliates excluded from this second subsample will tend to underreport their value added by overstating spending on inputs purchased within the firm (transfer pricing). These two subsamples thus drop affiliates for which, all else equal, if misreporting is present, the expected value of ε_{ijt} is positive (first subsample) or negative (second subsample). The resulting estimates appear in columns 1 through 4 in Table 10 and are very similar to those in Table 3, suggesting that value added misreporting is unlikely to explain our results.

Regarding R&D spending, the model in section 2 presumes actual and reported R&D expenditures coincide. Suppose instead that these differ. In particular, suppose that r_{ijt-1} is reported R&D spending, and that true R&D investment by affiliate j is $r_{ijt-1}^* \equiv r_{ijt-1} - x_{ijt-1}$, where x_{ijt-1} captures the difference between actual and reported R&D spending. We define an analogous pair of variables r_{i0t-1}^* and x_{i0t-1} corresponding to the parent. In this case, the error term in (16) becomes a function of both x_{ijt-1} and x_{i0t-1} . The mean independence condition in (17) will not hold if either deviation (x_{i0t-1} or x_{ijt-1}) is correlated with reported parent or affiliate R&D spending. Provided the true R&D investment of each firm site is positively correlated with its reporting error, it can be shown that such misreporting generates downward bias in estimates of μ_a , μ_p , and μ_{ap} computed using the procedure in section 4.1 (see Appendix A.15).⁶⁷ To obtain consistent estimates in the presence of misreporting, we re-estimate the model parameters replacing the estimator described in section 4.1 with a Generalized Method of Moments (GMM) estimator that relies on two policy instruments for identification.

Our first policy instrument aims to capture variation in affiliate R&D spending that is independent of local incentives to misreport. To build such an instrument, we combine information on U.S. parent innovation costs with measures of intellectual property rights in affiliate locations. Specifically, for affiliate j of firm i at t , this instrument interacts the Hall-Jorgenson user cost of R&D investment $UCRD_{it}$ prevailing in the firm- i U.S. parent state with an index of intellectual property rights $IPR_{n_{ijt}}$ in the affiliate host country at t .⁶⁸ The logic behind the interacted instrument is that an increase in parent R&D costs is expected to result in an optimal reallocation of R&D investment toward affiliates, particularly those in locations with strong patent rights. This is

⁶⁷A positive correlation between R&D spending and misreporting would likely arise, for example, in the presence of R&D subsidies, as these incentivize actual R&D investment and also over-reporting.

⁶⁸We obtain $UCRD$ from Wilson (2009); this variable has a mean of 1.192 (standard deviation 0.0494) during the sample period, and ranges from 1.035 to 1.338. IPR is from Ginarte and Park (1997) and Park (2008), has a mean value of 2.726 (standard deviation 1.050) during the sample period, and ranges between 0.20 to 4.67. For details on the construction of IPR , see Appendix A.3. This index is widely used; see, for example, Javorcik (2004b), McCalman (2004), Branstetter, Fisman, and Foley (2006), Qian (2007), and Bilir (2014).

consistent with the model described in section 2.6 provided that $UCRD_{it}$ enters as a determinant of the parent R&D cost shifter χ_{i0t}^r and $IPR_{n_{ijt}}$ affects affiliate profits multiplicatively as in Grossman and Lai (2004), in which case any impact of $IPR_{n_{ijt}}$ on affiliate revenue would be captured by the market-year effects $\kappa_{n_{ijt}}$ in (5). Estimates in Table A.2 align with this reasoning. The estimates show that, a) the user cost of R&D in a U.S. state is negatively correlated with R&D spending by U.S. parents located in that state, and b) is positively correlated with R&D spending by their foreign affiliates to an extent increasing in local intellectual property rights. These patterns are robust across specifications that include country-year and firm fixed effects (columns 2–3).

The validity of this instrument in the context of (16) is likely guaranteed by our inclusion of country-industry-year fixed effects. Importantly, any unobserved host-country characteristic that impacts firm incentives to misreport affiliate R&D (e.g. local tax rates) would not affect the validity of our instrument even if it is correlated with the strength of intellectual property protection $IPR_{n_{ijt}}$ across countries, as it would be controlled for by the set of country-sector-year fixed effects $\{\gamma_{n_{ijt}}\}$ included in (16). Notice this is true even if the impact of such an omitted variable on the incentives to misreport varies across sectors. By contrast, our instrument would be invalid in the presence of a host country characteristic that simultaneously a) affects affiliate incentives to misreport R&D, b) is correlated with $IPR_{n_{ijt}}$ in the affiliate host country, and c) affects affiliates differentially to an extent correlated with the user cost of R&D in the U.S. parent state.

Our second policy instrument is the measured local user cost of R&D, $UCRD_{it}$. There is a strong negative correlation between U.S. parent R&D investment and $UCRD_{it}$ in our sample, and also substantial spatial variation in the allocation of parents and parent R&D spending across U.S. states in the data (Zeile 2013). For our instrument to be valid, $UCRD_{it}$ must also be uncorrelated with the extent of parent R&D misreporting x_{i0t} . The validity of $UCRD_{it}$ thus requires a stronger assumption than does the validity of $UCRD_{it} \times IPR_{n_{ijt}}$: the variable $UCRD_{it}$ is a function of local R&D subsidies, which may encourage R&D overreporting by parents. However, the resulting correlation between $UCRD_{it}$ and x_{i0t-1} would cause downward bias in our estimates of the affiliate performance elasticity with respect to parent R&D spending (see Appendix A.15 for details).

To summarize, if R&D misreporting is present, we would expect: a) NLS estimates of the affiliate performance elasticity with respect to parent and affiliate R&D that are biased toward zero; and b) using the policy instruments described above, asymptotically unbiased GMM estimates of the affiliate performance elasticity with respect to affiliate R&D, and estimates of the elasticity with respect to parent R&D that are either asymptotically unbiased or biased toward zero.

Corresponding estimates for affiliates appear in columns 5 through 8 of Table 10. Comparing columns 5 and 6 with the analogous baseline estimates (Table 3, columns 2 and 4) reveals their qualitative similarity, but also indicates that the latter magnitudes are conservative. This is consistent with affiliates over-reporting innovation spending in countries with stronger misreporting incentives. In line with the hypotheses above, differences between our baseline (Table 3, column 2) and GMM estimates (Table 10, column 1) are also larger for the coefficient on affiliate R&D μ_a than for the coefficient on parent R&D μ_p . To address the bias that may arise if multinationals' U.S. headquarters location decisions are endogenous, columns 7 and 8 in Table 10 include U.S. state fixed effects. The resulting estimates are similar to those in columns 5 and 6; the positive coefficient

on the interaction between parent and affiliate R&D is the only coefficient sensitive to including these effects. GMM estimates for U.S. parents are included in Table A.3; the resulting estimate of μ_{0p} is larger than in the baseline specification, but retains its sign and significance. Taken together, our GMM estimates suggest that while misreporting may be present, the results regarding the within-firm impact of parent innovation on affiliate performance are robust to addressing this aspect of firm operations.

10 Conclusion

This paper examines the impact of parent innovation on affiliate performance in the multinational firm. Using data on the global activity of U.S.-based multinationals, we document that the spatial concentration of innovation is higher than that of production within these firms. We then estimate the parameters of a dynamic innovation model and find evidence that U.S. parent R&D investment systematically increases the performance of foreign affiliates within the same firm. These findings are consistent with the idea that knowledge is shared across locations within firm boundaries (Arrow 1975, Teece 1982), and with recent results emphasizing the transfer of intangible inputs as a key motive for plant integration (Atalay, Hortaçsu, and Syverson 2014).

Our analysis considers the intrafirm impact of a specific form of intangible input—the proprietary knowledge resulting from parent R&D investment—about which we observe uniquely detailed information in our data. From a policy perspective, the relevance of this input arises in part due to the prevalence of location-specific policies that aim to simulate private innovation investment. Our findings strongly suggest that such policies, where effective at increasing innovation by parents of multinational firms, create gains in within-firm productivity that are realized abroad.

In this respect, our results speak to a stated concern of the U.S. Congress regarding the allocation of gains from U.S. innovation subsidies. While further work is needed to determine the impact of aggregate U.S. R&D investment on welfare outcomes abroad, our results suggest a positive effect of U.S. innovation subsidies that is likely not internalized by policymakers. Adjusting the tax treatment of multinational firms offers one potential remedy, but deeper international coordination of innovation policy may be an effective alternative. In particular, building on the global patent protection treaties of the last century, our estimates suggests that for industries dominated by offshoring and multinational production, further efficiency gains could result from the international coordination of innovation subsidies.

REFERENCES

- [1] **Akerberg, Daniel A., Kevin Caves, and Garth Frazer.** (2015). “Identification Properties of Recent Production Function Estimators.” *Econometrica*, vol. 83, no. 6, pp. 2411–2451.
- [2] **Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple.** (2014). “Innovation and Production in the Global Economy.” National Bureau of Economic Research Working Paper 18792.
- [3] **Arnold, Jens Matthias, and Beata Javorcik.** (2009). “Gifted Kids Or Pushy Parents? Foreign Direct Investment and Plant Productivity in Indonesia.” *Journal of International Economics*, vol. 79, no. 1, pp. 42–53.
- [4] **Arrow, Kenneth J.** (1962). “Economic Welfare and the Allocation of Resources for Invention,” in *The Rate and Direction of Inventive Activity: Economic and Social Factors*, National Bureau of Economic Research.
- [5] **Arrow, Kenneth J.** (1969). “Classificatory Notes on the Production and Transmission of Technological Knowledge.” *American Economic Review, Papers and Proceedings*, vol. 59, no. 2, pp. 29–35.
- [6] **Arrow, Kenneth J.** (1975). “Vertical Integration and Communication.” *Bell Journal of Economics*, vol. 6, no. 1, pp. 173–183.
- [7] **Atalay, Engin, Ali Hortaçsu, and Chad Syverson.** (2014). “Vertical Integration and Input Flows.” *American Economic Review*, vol. 104, no. 4, pp. 1120–1148.
- [8] **Atkeson, Andrew, and Ariel Burstein.** (2015). “The Aggregate Implications of Innovation Policy.” University of California mimeo.
- [9] **Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu.** (2008). “R&D Investment, Export, and the Evolution of Firm Productivity.” *American Economic Review, Papers and Proceedings*, vol. 98, no. 2, pp. 451–456.
- [10] **Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu.** (2011). “R&D Investment, Exporting, and Productivity Dynamics.” *American Economic Review*, vol. 101, no. 4, pp. 1312–1344.
- [11] **Balat, Jorge, Irene Brambilla, and Yuya Sasaki.** (2016). “Heterogeneous Firms: Skilled-Labor Productivity and Export Destinations.” Johns Hopkins University mimeo.
- [12] **Bernard, Andrew J., Bradford Jensen, and Peter Schott.** (2006). “Transfer Pricing by U.S.-Based Multinational Firms.” Yale University mimeo.
- [13] **Bilir, L. Kamran.** (2014). “Patent Laws, Product Lifecycle Lengths, and Multinational Activity.” *American Economic Review*, vol. 104, no. 7, pp. 1979–2013.
- [14] **Bøler, Esther A.** (2015). “Technology-skill Complementarity in a Globalized World.” University of Oslo mimeo.
- [15] **Bøler, Esther A., Andreas Moxnes, and Karen H. Ulltveit-Moe.** (2015). “R&D, International Sourcing, and the Joint Impact on Firm Performance.” *American Economic Review*, vol. 105, no. 12, pp. 3704–3739.
- [16] **Branstetter, Lee G., Raymond Fisman, and C. Fritz Foley.** (2006). “Do Stronger Intellectual Property Rights Increase International Technology Transfer? Empirical Evidence from U.S. Firm-level Panel Data.” *Quarterly Journal of Economics*, vol. 121, no. 1, pp. 321–349.
- [17] **Cohen, Wesley M., and Steven Klepper.** (1996). “Firm Size and the Nature of Innovation within Industries: The Case of Process and Product R&D.” *Review of Economics and Statistics*, vol. 78, no. 2, pp. 232–243.
- [18] **Cohen, Wesley M., and David A. Levinthal.** (1989). “Innovation and Learning.” *Economic Journal*, vol. 99, pp. 569–596.
- [19] **Criscuolo, Chiara, Jonathan E. Haskel, and Matthew Slaughter.** (2010). “Global Engagement and the Innovation Activities of Firms.” *International Journal of Industrial Organization*, vol. 28, no. 2, pp. 191–202.

- [20] **De Loecker, Jan.** (2011). “Product Differentiation, Multi-Product Firms and Estimating the Impact of Trade Liberalization on Productivity.” *Econometrica*, vol. 79, no. 5, pp. 1407–1451.
- [21] **De Loecker, Jan, and Frederic Warzynski.** (2012). “Markups and Firm Level Export Status.” *American Economic Review*, vol. 102, no. 6, pp. 2437–2471.
- [22] **De Loecker, Jan, Pienelopi K. Goldberg, Amit K. Khandelwal and Nina Pavcnik.** (2016). “Prices, Markups and Trade Reform.” *Econometrica*, vol. 84, no. 2, pp. 445–510.
- [23] **Dickstein, Michael, and Eduardo Morales.** (2015). “What Do Exporters Know?” Princeton University mimeo.
- [24] **Dhingra, Swati.** (2013). “Trading Away Wide Brands for Cheap Brands.” *American Economic Review*, vol. 103, no. 6, pp. 2554–2584.
- [25] **Doraszelski, Ulrich, and Jordi Jaumandreu.** (2013). “R&D and Productivity: Estimating Endogenous Productivity.” *Review of Economic Studies*, vol. 80, pp. 1338–1383.
- [26] **Doraszelski, Ulrich, and Jordi Jaumandreu.** (2016). “Measuring the Bias of Technological Change.” University of Pennsylvania mimeo.
- [27] **Dunne, Timothy, Mark J. Roberts and Larry Samuelson.** (1989). “The Growth and Failure of U. S. Manufacturing Plants.” *Quarterly Journal of Economics*, vol. 104, no. 4, pp. 671–698.
- [28] **Evans, David.** (1987). “Tests of Alternative Theories of Firm Growth.” *Journal of Political Economy*, vol. 95, no. 4, pp. 657–674.
- [29] **Fillat, José L., and Stefania Garetto.** (2015). “Risk, Returns, and Multinational Production.” *Quarterly Journal of Economics*, vol. 130, no. 4, pp. 2027–2073.
- [30] **Foster, Lucia, John Haltiwanger, and Chad Syverson.** (2008). “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?” *American Economic Review*, vol. 98, no. 1, pp. 394–495.
- [31] **Frisch, Ragnar, and Frederick V. Waugh.** (1933). “Time Regressions as Compared with Individual Trends.” *Econometrica*, vol. 1, no. 4, pp. 387–401.
- [32] **Gandhi, Amit, Salvador Navarro, and David Rivers.** (2016). “On the Identification of Production Functions: How Heterogeneous is Productivity?” University of Wisconsin-Madison mimeo.
- [33] **Garetto, Stefania, Lindsay Oldenski, and Natalia Ramondo.** (2016). “Life-Cycle Dynamics and the Expansion Strategies of U.S. Multinational Firms.” Boston University mimeo.
- [34] **Giles, David E. A.** (1984). “Instrumental Variables Regressions Involving Seasonal Data.” *Economics Letters*, vol. 14, no. 4, pp. 339–343.
- [35] **Ginarte, Juan C., and Walter G. Park.** (1997). “Determinants of Patent Rights: A Cross-National Study.” *Research Policy*, vol. 26, no. 3, pp. 283–301.
- [36] **Giraud, Xavier.** (2013). “Proximity and Investment: Evidence from Plant-level Data.” *The Quarterly Journal of Economics*, vol. 128, no. 2, pp. 861–915.
- [37] **Government Accountability Office.** (2009). “Tax Policy: The Research Tax Credit’s Design and Administration Can Be Improved,” Report to the Committee on Finance, U.S. Senate, November.
- [38] **Gravelle, Jane G.** (2015). “Tax Havens: International Tax Avoidance and Tax Evasion.” *Congressional Research Service Report*, no. 7-5700, pp. 605–627.
- [39] **Griliches, Zvi.** (1979). “Issues in Assessing the Contribution of Research and Development to Productivity Growth.” *Bell Journal of Economics*, vol. 10, no. 1, pp. 92–116.
- [40] **Griliches, Zvi, and Philippe Mairesse.** (1998). “Production Functions: The Search for Identification,” In *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*. New York: Cambridge University Press, 169–203.
- [41] **Grossman, Gene M., and Edwin L. C. Lai.** (2004). “International Protection of Intellectual Property.” *American Economic Review*, vol. 94, no. 5, pp. 1635–1653.

- [42] **Guadalupe, María, Olga Kuzmina, and Catherine Thomas.** (2012). “Innovation and Foreign Ownership.” *American Economic Review*, vol. 102, no. 7, pp. 3594–3627.
- [43] **Gumpert, Anna.** (2015). “The Organization of Knowledge in Multinational Firms.” University of Munich mimeo.
- [44] **Gumpert, Anna, Andreas Moxnes, Natalia Ramondo, and Felix Tintelnot.** (2016). “Exporters’ and Multinational Firms’ Life-cycle Dynamics.” University of Chicago mimeo.
- [45] **Hall, Robert E., and Dale Jorgenson.** (1967). “Tax Policy and Investment Behavior.” *American Economic Review*, vol. 57, no. 3, pp. 391–414.
- [46] **Haltiwanger, John, Ron S. Jarmin, and Javier Miranda.** (2013). “Who Creates Jobs? Small Versus Large Versus Young.” *The Review of Economics and Statistics*, vol. 95, no. 2, pp. 541–559.
- [47] **Head, Keith, and Thierry Mayer.** (2014). “Gravity Equations: Workhorse, Toolkit, and Cookbook.” in Gopinath, Gita, Elhanan Helpman, and Kenneth Rogoff (editors), *Handbook of International Economics*, Elsevier, vol. 4, pp. 131–195.
- [48] **Head, Keith, and Thierry Mayer.** (2016). “Brands in Motion: How Frictions Shape Multinational Production.” University of British Columbia mimeo.
- [49] **Helpman, Elhanan.** (1984). “A Simple Theory of International Trade with Multinational Corporations.” *Journal of Political Economy*, vol. 92, no. 3, pp. 451–471.
- [50] **Hines, James R., Jr.** (1995). “Taxes, Technology Transfer, and R&D by Multinational Firms,” in Hines, James R., James R. Hines, Jr., and R. Glenn Hubbard (editors), *Taxing Multinational Corporations*, University of Chicago Press.
- [51] **Hines, James R., Jr.** (1997). “Tax Policy and the Activities of Multinational Corporations,” in: Auerbach, A.J. (Ed.) *Fiscal Policy: Lessons from Economic Research*. MIT Press, Cambridge: 401–445.
- [52] **Hines, James R., Jr., and Eric M. Rice.** (1994). “Fiscal Paradise: Foreign Tax Havens and American Business.” *Quarterly Journal of Economics*, vol 109, pp. 149–181.
- [53] **Holmes, Thomas J.** (2011). “The Diffusion of Wal-Mart and Economies of Density.” *Econometrica*, vol. 79, no. 1, pp. 253–302.
- [54] **Holmes, Thomas J., Ellen R. McGrattan, and Edward C. Prescott.** (2015). “Quid Pro Quo: Technology Capital Transfers for Market Access in China.” *Review of Economic Studies*, vol. 82, no. 3, pp. 1154–1193.
- [55] **Huergo, Elena, and Jordi Jaumandreu.** (2004). “Firms Age, Process Innovation and Productivity Growth.” *International Journal of Industrial Organization*, vol. 22, pp. 541–559.
- [56] **Irrazabal, Alfonso, Andreas Moxnes, and Luca David Opromolla.** (2013). “The Margins of Multinational Production and the Role of Intrafirm Trade.” *Journal of Political Economy*, vol. 121, no. 1, pp. 74–126.
- [57] **Jaumandreu, Jordi, and Heng Yin.** (2014). “Cost and Product Advantages: A Firm-level Model for Chinese Exports and Industry Growth.” Boston University mimeo.
- [58] **Javorcik, Beata.** (2004a). “Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers Through Backward Linkages.” *American Economic Review*, vol. 94, no. 3, pp. 605–627.
- [59] **Javorcik, Beata.** (2004b). “The Composition of Foreign Direct Investment and Protection of Intellectual Property Rights: Evidence from Transition Economies.” *European Economic Review*, no. 48, pp. 39–62.
- [60] **Javorcik, Beata, and Steven Poelhekke.** (2016). “Former Foreign Affiliates: Cast Out and Outperformed?” *forthcoming, Journal of the European Economic Association*.
- [61] **Kasahara, Hiroyuki, Paul Schrimpf, and Michio Suzuki.** (2015). “Identification and Estimation of Production Function with Unobserved Heterogeneity.” University of British Columbia mimeo.

- [62] **Keller, Wolfgang.** (2002). “Geographic Localization of International Technology Diffusion.” *American Economic Review*, vol. 92, no. 1, pp. 120–142.
- [63] **Keller, Wolfgang.** (2004). “International Technology Diffusion.” *Journal of Economic Literature*, vol. 42, no. 3, pp. 752–782.
- [64] **Keller, Wolfgang, and Stephen R. Yeaple.** (2013). “The Gravity of Knowledge.” *American Economic Review*, vol. 103, no. 4, pp. 1414–1444.
- [65] **Klette, Tor Jakob, and Samuel Kortum.** (2004). “Innovating Firms and Aggregate Innovation.” *Journal of Political Economy*, vol. 112, no. 5, pp. 986–1018.
- [66] **Levinsohn, James, and Amil Petrin.** (2003). “Estimating Production Functions Using Inputs to Control for Unobservables.” *Review of Economic Studies*, vol. 70, pp. 317–341.
- [67] **Lovell, Michael C.** (1963). “Seasonal Adjustment of Economic Time Series and Multiple Regression Analysis.” *Journal of the American Statistical Association*, vol. 58, no. 304, pp. 993–1010.
- [68] **Markusen, James R.** (1984). “Multinationals, Multi-Plant Economies, and the Gains from Trade.” *Journal of International Economics*, vol. 16, no. 3, pp. 205–226.
- [69] **Mataloni, Raymond J. Jr., and Lee Goldberg.** (1994). “Gross Product of U.S. Multinational Companies, 1977–2001.” *Survey of Current Business*, vol. 70, (February), pp. 42–63.
- [70] **McCalman, Phillip.** (2004). “Foreign Direct Investment and Intellectual Property Rights: Evidence from Hollywood’s Global Distribution of Movies and Videos.” *Journal of International Economics*, vol. 62, pp. 107–123.
- [71] **McGrattan, Ellen R.** (2012). “Transition to FDI Openness: Reconciling Theory and Evidence.” *Review of Economic Dynamics*, vol. 15, no. 4, pp. 437–458.
- [72] **McGrattan, Ellen R., and Edward C. Prescott.** (2010). “Technology Capital and the U.S. Current Account.” *American Economic Review*, vol. 100, no. 4, pp. 1493–1522.
- [73] **Morales, Eduardo, Gloria Sheu, and Andres Zahler.** (2015). “Extended Gravity.” Princeton University mimeo.
- [74] **Nadiri, M. Ishaq, and Ingmar Prucha.** (1996). “Estimation of the Depreciation Rate of Physical and R&D Capital in the U.S. Total Manufacturing Sector.” *Economic Inquiry*, vol. 34, pp. 43–56.
- [75] **National Science Board.** (2014). *Science and Engineering Indicators 2014*. Arlington, VA: National Science Foundation (NSB 14-01).
- [76] **Olley, Steven G., and Ariel Pakes.** (1996). “The Dynamics of Productivity in the Telecommunications Equipment Industry.” *Econometrica*, vol. 64, no. 6, pp. 1263–1295.
- [77] **Park, Walter G.** (2008). “International Patent Protection: 1960–2005.” *Research Policy*, vol. 37, no. 4, pp. 761–766.
- [78] **Qian, Yi.** (2007). “Do National Patent Laws Stimulate Domestic Innovation In A Global Patenting Environment?” *Review of Economics and Statistics*, vol. 89, no. 3, pp. 436–453.
- [79] **Ramondo, Natalia, Veronica Rappoport, and Kim Ruhl.** (2016). “Intra-firm Trade and Vertical Fragmentation in U.S. Multinational Corporations.” *Journal of International Economics*, vol. 98, pp. 51–59.
- [80] **Ramondo, Natalia, and Andrés Rodríguez-Clare.** (2013). “Trade, Multinational Production, and the Gains from Openness.” *Journal of Political Economy*, vol. 121, no. 2, pp. 273–322.
- [81] **Rao, Nirupama S.** (2016). “Do Tax Credits Stimulate R&D Spending? The Effect of the R&D Tax Credit in its First Decade.” *Journal of Public Economics, forthcoming*.
- [82] **Roberts, Mark J., Daniel Yi Xu, Xiaoyan Fan and Shengxing Zhang.** (2011). “A Structural Model of Demand, Cost, and Export Market Selection for Chinese Footwear Producers.” *NBER Working Paper*, 17725.

- [83] **Rodríguez-Clare, Andrés.** (1996). “Multinationals, Linkages and Economic Development.” *American Economic Review*, vol. 86, no. 4, pp. 852–873.
- [84] **Schankerman, Mark.** (1981). “The Effects of Double-Counting and Expensing on the Measured Returns to R&D.” *Review of Economics and Statistics*, vol. 63, no. 3, pp. 454–459.
- [85] **Teece, David J.** (1982). “Towards an Economic Theory of the Multiproduct Firm.” *Journal of Economic Behavior and Organization*, vol. 3, no. 1, pp. 39–63.
- [86] **Tintelnot, Felix.** (2016). “Global Production with Export Platforms.” *Quarterly Journal of Economics*, forthcoming.
- [87] **United Nations Conference on Trade and Development.** (2005). *World Investment Report 2005: Transnational Corporations and the Internationalization of R&D*. New York: United Nations.
- [88] **U.S. Bureau of Economic Analysis.** (2008). “U.S. Direct Investment Abroad: 2004 Final Benchmark Data.” Washington, DC: U.S. Government Printing Office, November 2008.
- [89] **U.S. Senate Permanent Subcommittee on Investigations.** (2013). “Offshore Profit Shifting and the U.S. Tax Code.” U.S. Senate Hearing, May 21.
- [90] **Wilson, Daniel J.** (2009). “Beggar Thy Neighbor? The In-State, Out-of-State, and Aggregate Effects of R&D Tax Credits.” *Review of Economics and Statistics*, vol. 91, no. 2, pp. 431–436.
- [91] **Zeile, William J.** (2013). “State-Level R&D by Multinational Companies Results From an Interagency Data Link Project.” *Survey of Current Business*, vol. 93, no. 9, pp. 286–293.

Table 1: Descriptive Statistics, R&D Allocation in the Multinational Firm

Variable	<u>Firm-Year Level</u>	
	Mean	Std. Dev.
Percentage of Firms with Positive Parent R&D Expenditure	92.5%	
Percentage of Firms with Positive Affiliate R&D Expenditure	47.8%	
Percentage of Affiliates per Firm with Positive R&D Expenditure	24.1%	34.7%
Affiliate Share in Total Firm R&D Expenditure	8.5%	19.4%
Affiliate Share in Total Firm Value Added	16.7%	51.8%
Affiliate Share in Total Firm Sales	28.8%	20.4%
Affiliate Share in Total Firm Employment	26.6%	22.6%

All variables are from the 1994 Bureau of Economic Analysis Benchmark Survey of U.S. Direct Investment Abroad; the 1994 survey is unusually comprehensive in its coverage of U.S. multinational firms' activity abroad. These data are a cross section covering U.S.-based multinational firms operating in the Computers and Office Equipment industry (SIC 357) including U.S. parent sites and foreign manufacturing affiliates.

Table 2: Summary Statistics, 1989-2008

Variable	<u>Parent-Year Level</u>		<u>Affiliate-Year Level</u>	
	Mean	Std. Dev.	Mean	Std. Dev.
Value Added (thousands \$US)	1,140,000	3,090,000	158,000	442,000
Sales (thousands \$US)	3,710,000	9,060,000	664,000	1,490,000
Value of Plant and Equipment (thousands \$US)	666,000	2,170,000	121,000	372,000
R&D Expenditure (thousands \$US)	300,000	856,000	7,660	35,100
Number of Workers	9,170	22,800	1,500	3,150
Share of R&D Workers	17.2%	11.6%	15.7%	19.8%
R&D Expenditure/Sales	7.5%	11.5%	1.1%	2.7%
Imports from Parent (\$US)			81,900	262,000
Royalty Payments to Parent (\$US)			42,100	222,000
Observations	536	536	4,194	4,194

Notes: All variables are from the Bureau of Economic Analysis Survey of U.S. Direct Investment Abroad. These panel data span 1989-2008 and cover U.S.-based multinational firms operating in the Computers and Office Equipment industry (SIC 357) including U.S. parent sites and foreign manufacturing affiliates.

Table 3: Baseline Estimates

	NLS			
	(1)	(2)	(3)	(4)
Persistence	0.7915 ^a (0.0095)	0.7721 ^a (0.0098)	0.7722 ^a (0.0098)	0.7681 ^a (0.0098)
Affiliate R&D	0.0064 ^a (0.0013)	0.0042 ^a (0.0013)	0.0043 ^a (0.0014)	-0.0272 ^a (0.0070)
Parent R&D		0.0164 ^a (0.0022)	0.0166 ^a (0.0023)	0.0103 ^a (0.0026)
Other Affiliates' R&D			-0.0003 (0.0013)	
Affiliate R&D \times Parent R&D				0.0025 ^a (0.0005)
Labor Elasticity	0.4682 (0.0646)	0.4682 (0.0646)	0.4682 (0.0646)	0.4682 (0.0646)
Capital Elasticity	0.2526 (0.0572)	0.2488 (0.0536)	0.2487 (0.0536)	0.2447 (0.0517)
Observations	4,194	4,194	4,194	4,194

Notes: *a* denotes 1% significance, *b* denotes 5% significance, *c* denotes 10% significance. All columns report Nonlinear Least Squares estimates of the parameters in (9) and several variants thereof. All columns include market-year fixed effects. Standard errors are reported in parenthesis. Persistence corresponds to estimates of ρ . Affiliate R&D, Parent R&D, Other Affiliates' R&D and Affiliate R&D \times Parent R&D estimates capture the elasticity of period t performance with respect to the period $t - 1$ value of the corresponding covariate. Labor Elasticity is the average value of $\beta_l + \beta_{ll}2l_{ijt} + \beta_{lk}k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk}2k_{ijt} + \beta_{lk}l_{ijt}$. The standard deviation for each of these input elasticities appears in parentheses below its mean. Measures of labor, capital, value added, and R&D expenditure are from the Bureau of Economic Analysis Survey of U.S. Direct Investment Abroad.

Table 4: Estimates with Affiliate Royalty Payments to U.S. Parent

	(1)	(2)	(3)	(4)
Persistence	0.7463 ^a (0.0101)	0.7445 ^a (0.0101)	0.7411 ^a (0.0102)	0.7398 ^a (0.0102)
Affiliate R&D	0.0027 ^b (0.0013)	-0.0203 ^a (0.0060)	0.0024 ^c (0.0013)	-0.0173 ^b (0.0073)
Parent R&D	0.0128 ^a (0.0022)	0.0085 ^a (0.0025)	0.0072 ^a (0.0026)	0.0038 (0.0029)
Affiliate R&D \times Parent R&D		0.0018 ^a (0.0005)		0.0016 ^a (0.0006)
Parent Royalties	0.0112 ^a (0.0012)	0.0107 ^a (0.0012)	-0.0146 ^a (0.0062)	-0.0129 ^c (0.0064)
Parent R&D \times Parent Royalties			0.0020 ^a (0.0005)	0.0019 ^a (0.0005)
Affiliate R&D \times Parent R&D \times Parent Royalties				-0.0000 (0.0000)
Labor Elasticity	0.4682 (0.0646)	0.4682 (0.0646)	0.4682 (0.0646)	0.4682 (0.0646)
Capital Elasticity	0.2370 (0.0499)	0.2350 (0.0488)	0.2345 (0.0487)	0.2295 (0.0470)
Observations	4,194	4,194	4,194	4,194

Notes: *a* denotes 1% significance, *b* denotes 5% significance, *c* denotes 10% significance. This table reports Nonlinear Least Squares estimates corresponding to variants of equation (9) that incorporate observed affiliate royalties paid to the U.S. parent (Parent Royalties) for the use of proprietary technology. All columns include market-year fixed effects. Standard errors are reported in parenthesis. Persistence corresponds to estimates of ρ . Labor Elasticity is the average value of $\beta_l + \beta_{ll}2l_{ijt} + \beta_{lk}k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk}2k_{ijt} + \beta_{lk}l_{ijt}$. The standard deviation for each of these input elasticities appears in parentheses below its mean. All other estimates capture the elasticity of period t performance with respect to the $t - 1$ value of the corresponding covariate. Measures of labor, capital, value added, R&D expenditure, and affiliate royalty payments to the parent are from the Bureau of Economic Analysis Survey of U.S. Direct Investment Abroad.

Table 5: Flexible Functions of R&D Investment

	Additional R&D Lags			Allowing for R&D Dummies				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Persistence	0.8096 ^a (0.0106)	0.8067 ^a (0.0106)	0.8253 ^a (0.0114)	0.8230 ^a (0.0114)	0.7880 ^a (0.0094)	0.7682 ^a (0.0097)	0.7682 ^a (0.0097)	0.7398 ^a (0.0097)
Affiliate R&D ($t-1$)					0.0230 ^a (0.0034)	0.0189 ^a (0.0034)	0.0190 ^a (0.0034)	-0.0127 (0.0091)
Parent R&D ($t-1$)						0.0170 ^a (0.0022)	0.0171 ^a (0.0024)	0.0117 ^a (0.0026)
Other Affiliates' R&D ($t-1$)							-0.0002 (0.0013)	
Affiliate R&D \times Parent R&D ($t-1$)								0.0022 ^a (0.0006)
Affiliate R&D ($t-2$)	0.0052 ^a (0.0014)	-0.0239 ^a (0.0081)						
Parent R&D ($t-2$)	0.0154 ^a (0.0027)	0.0084 ^a (0.0033)						
Affiliate R&D \times Parent R&D ($t-2$)		0.0023 ^a (0.0006)						
Affiliate R&D ($t-3$)			0.0051 ^a (0.0015)	-0.0140 (0.0089)				
Parent R&D ($t-3$)			0.0162 ^a (0.0030)	0.0114 ^a (0.0037)				
Affiliate R&D \times Parent R&D ($t-3$)				0.0015 ^b (0.0007)				
Labor Elasticity	0.4684 (0.0625)	0.4684 (0.0625)	0.4706 (0.0613)	0.4706 (0.0613)	0.4682 (0.0646)	0.4682 (0.0646)	0.4682 (0.0646)	0.4682 (0.0646)
Capital Elasticity	0.2427 (0.0506)	0.2397 (0.0496)	0.2311 (0.0473)	0.2297 (0.0466)	0.2381 (0.0508)	0.2344 (0.0478)	0.2344 (0.0478)	0.2334 (0.0473)
Observations	3,631	3,631	3,086	3,086	4,194	4,194	4,194	4,194

Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. This table reports Nonlinear Least Squares estimates corresponding to several variants of (9). Columns (1) to (4) add further lags of parent and affiliate R&D investment; columns (5) through (8) include a second set of market-year fixed effects that are interacted with a dummy variable capturing R&D spending by the affiliate. All columns include market-year fixed effects. Standard errors are reported in parentheses. Persistence corresponds to estimates of ρ . Labor Elasticity is the average value of $\beta_l + \beta_{ll}2l_{ijt} + \beta_{lk}k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk}2k_{ijt} + \beta_{kl}l_{ijt}$. The standard deviation for each of these input elasticities appears in parentheses below its mean. All other estimates capture the elasticity of period t performance with respect to the $t-1$ value of the corresponding covariate. Measures of labor, capital, value added, and R&D expenditure are from the Bureau of Economic Analysis Survey of U.S. Direct Investment Abroad.

Table 6: Additional Performance Lags, Correlated Performance Shocks, and Labor Adjustment

	Additional Productivity Lags				Firm Shocks		Non-R&D Labor	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Persistence ($t - 1$)	0.7206 ^a (0.0155)	0.7162 ^a (0.0155)	0.6535 ^a (0.0222)	0.6497 ^a (0.0222)	0.7721 ^a (0.0255)	0.7681 ^a (0.0255)	0.7739 ^a (0.0101)	0.7702 ^a (0.0101)
Persistence ($t - 2$)	0.1267 ^a (0.0151)	0.1256 ^a (0.0151)	0.1475 ^a (0.0267)	0.1443 ^a (0.0266)				
Persistence ($t - 3$)			0.0881 ^a (0.0211)	0.0882 ^a (0.0211)				
Affiliate R&D	0.0013 (0.0013)	-0.0255 ^a (0.0070)	0.0010 (0.0016)	-0.0237 ^a (0.0090)	0.0042 ^b (0.0017)	-0.0272 ^a (0.0097)	0.0043 ^a (0.0013)	-0.0251 ^a (0.0068)
Parent R&D	0.0145 ^a (0.0022)	0.0094 ^a (0.0026)	0.0136 ^a (0.0027)	0.0098 ^a (0.0031)	0.0164 ^a (0.0039)	0.0103 ^b (0.0051)	0.0136 ^a (0.0021)	0.0081 ^a (0.0025)
Affiliate R&D \times Parent R&D		0.0021 ^a (0.0005)		0.0019 ^a (0.0007)		0.0025 ^a (0.0008)		0.0023 ^a (0.0005)
Labor Elasticity	0.4689 (0.0620)	0.4689 (0.0620)	0.4761 (0.0589)	0.4761 (0.0589)	0.4682 (0.0646)	0.4682 (0.0646)	0.4709 (0.0649)	0.4709 (0.0649)
Capital Elasticity	0.2390 (0.0407)	0.2364 (0.0402)	0.2503 (0.0383)	0.2487 (0.0384)	0.2488 (0.0536)	0.2447 (0.0517)	0.2620 (0.0565)	0.2585 (0.0550)
Observations	3,442	3,442	1,984	1,984	4,194	4,194	4,194	4,194

Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. This table reports Nonlinear Least Squares estimates corresponding to several variants of (9). Columns (1) and (2) control for a second productivity lag at $t - 2$; columns (3) and (4) also include productivity at $t - 3$; columns (5) and (6) cluster standard errors at the multinational firm-year level. All columns include market-year fixed effects. Standard errors are reported in parentheses. Persistence corresponds to estimates of ρ . Labor Elasticity is the average value of $\beta_l + \beta_{l2l_{i,t}} + \beta_{lk}k_{i,t}$; Capital Elasticity is the average value of $\beta_k + \beta_{k2k_{i,t}} + \beta_{lk}l_{i,t}$. The standard deviation for each of these input elasticities appears in parentheses below its mean. All other estimates capture the elasticity of period t performance with respect to the $t - 1$ value of the corresponding covariate. Measures of labor, capital, value added, and R&D expenditure are from the Bureau of Economic Analysis Survey of U.S. Direct Investment Abroad.

Table 7: Industry Restriction and Markup Heterogeneity

	Parent Sector		Heterogeneous Markups				<i>Distribution of ν_{2t}</i>
	(1)	(2)	(3)	(4)	(5)	(6)	
Persistence	0.8271 ^a (0.0253)	0.8266 ^a (0.0247)	0.7982 ^a (0.0089)	0.7808 ^a (0.0092)	0.7808 ^a (0.0093)	0.7833 ^a (0.0093)	minimum 2.07
Affiliate R&D	0.0017 (0.0046)	-0.1015 ^a (0.0261)	0.0021 ^a (0.0004)	0.0015 ^a (0.0004)	0.0015 ^a (0.0004)	-0.0026 ^c (0.0014)	percentile 25 2.63
Parent R&D	0.0166 ^a (0.0075)	-0.0002 (0.0085)	0.0049 ^a (0.0007)	0.0049 ^a (0.0007)	0.0049 ^a (0.0000)	0.0041 ^a (0.0008)	median 2.84
Other Affiliates' R&D					0.0000 (0.0004)		percentile 75 3.33
Affiliate R&D \times Parent R&D		0.0077 ^a (0.0019)			0.0003 ^a (0.0001)		maximum 3.72
Labor Elasticity	0.4249 (0.0935)	0.4249 (0.0935)	0.4702 (0.0831)	0.4702 (0.0831)	0.4702 (0.0831)	0.4702 (0.0831)	average 2.91
Capital Elasticity	0.1437 (0.1167)	0.1212 (0.1058)	0.1986 (0.0503)	0.1974 (0.0488)	0.1975 (0.0488)	0.1961 (0.0483)	std. dev. 0.40
Observations	582	582	4,161	4,161	4,161	4,161	

Notes: *a* denotes 1% significance, *b* denotes 5% significance, *c* denotes 10% significance. This table reports Nonlinear Least Squares estimates corresponding to several variants of (9). Columns (1) and (2) restrict the sample to include only affiliates operating in the same three-digit industry as the U.S. parent (computers, SIC 357); columns (3) through (6) allow for heterogeneous affiliate markups across periods *t*; and the distribution of ν_2 appears in the last column. All columns include market-year fixed effects. Standard errors are reported in parenthesis. Persistence corresponds to estimates of ρ . Labor Elasticity is the average value of $\beta_l + \beta_{ll}2k_{ijt} + \beta_{lk}k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk}2k_{ijt} + \beta_{lk}l_{ijt}$. The standard deviation for each of these input elasticities appears in parentheses below its mean. All other estimates capture the elasticity of period *t* performance with respect to the *t* - 1 value of the corresponding covariate. Measures of labor, capital, value added, and R&D expenditure are from the Bureau of Economic Analysis Survey of U.S. Direct Investment Abroad.

Table 8: Other Sectors

<i>Panel A: Other 3-digit SIC Sectors</i>								
	<i>Drugs</i>				<i>Motor Vehicles and Motor Vehicle Equipment</i>			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Persistence	0.7910 ^a (0.0072)	0.7879 ^a (0.0073)	0.7879 ^a (0.0073)	0.7860 ^a (0.0073)	0.6755 ^a (0.0067)	0.6696 ^a (0.0068)	0.6693 ^a (0.0068)	0.6696 ^a (0.0068)
Affiliate R&D	0.0112 ^a (0.0009)	0.0108 ^a (0.0009)	0.0107 ^a (0.0009)	-0.0029 (0.0064)	0.0035 ^a (0.0008)	0.0028 ^a (0.0008)	0.0026 ^a (0.0008)	-0.0004 (0.0038)
Parent R&D		0.0066 ^a (0.0020)	0.0055 ^a (0.0023)	0.0026 ^a (0.0027)		0.0087 ^a (0.0012)	0.0078 ^a (0.0014)	0.0081 ^a (0.0014)
Other Affiliates' R&D			0.0011 (0.0013)				0.0012 (0.0009)	
Affiliate R&D × Parent R&D				0.0010 ^a (0.0005)				0.0003 (0.0003)
Labor Elasticity	0.4298 (0.0453)	0.4298 (0.0453)	0.4298 (0.0453)	0.4298 (0.0453)	0.5698 (0.0808)	0.5698 (0.0808)	0.5698 (0.0808)	0.5698 (0.0808)
Capital Elasticity	0.1782 (0.0392)	0.1779 (0.0386)	0.1779 (0.0386)	0.1778 (0.0247)	0.2129 (0.0673)	0.2092 (0.0641)	0.2093 (0.0641)	0.2084 (0.0637)
Observations	7,285	7,285	7,285	7,285	9,953	9,953	9,953	9,953
<i>Panel B: 2-digit SIC Sectors</i>								
	<i>Industrial and Commercial Machinery & Computer Equipment</i>				<i>Chemicals and Allied Products</i>			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Persistence	0.7188 ^a (0.0092)	0.7120 ^a (0.0093)	0.7106 ^a (0.0094)	0.7113 ^a (0.0093)	0.7469 ^a (0.0060)	0.7436 ^a (0.0061)	0.7427 ^a (0.0061)	0.7372 ^a (0.0061)
Affiliate R&D	0.0049 ^a (0.0011)	0.0049 ^a (0.0011)	0.0048 ^a (0.0011)	-0.0053 (0.0045)	0.0102 ^a (0.0007)	0.0096 ^a (0.0007)	0.0092 ^a (0.0007)	-0.0133 ^a (0.0031)
Parent R&D		0.0101 ^a (0.0019)	0.0092 ^a (0.0020)	0.0064 ^a (0.0024)		0.0040 ^a (0.0010)	0.0021 ^c (0.0012)	-0.0005 (0.0012)
Other Affiliates' R&D			0.0013 (0.0011)				0.0027 ^a (0.0008)	
Affiliate R&D × Parent R&D				0.0009 ^a (0.0003)				0.0019 ^a (0.0002)
Labor Elasticity	0.5441 (0.0779)	0.5441 (0.0779)	0.5441 (0.0779)	0.5441 (0.0779)	0.4421 (0.0534)	0.4421 (0.0534)	0.4421 (0.0534)	0.4421 (0.0534)
Capital Elasticity	0.1729 (0.0319)	0.1682 (0.0283)	0.1688 (0.0283)	0.1674 (0.0279)	0.1611 (0.0306)	0.1604 (0.0301)	0.1604 (0.0298)	0.1609 (0.0293)
Observations	5,016	5,016	5,016	5,016	10,681	10,681	10,681	10,681

Notes: *a* denotes 1% significance, *b* denotes 5% significance, *c* denotes 10% significance. All columns report Nonlinear Least Squares estimates corresponding to (16) as specified in section 2 and several variants thereof for firms in additional industries. In Panel A, columns (1) to (4) report estimates for affiliates whose parent is in SIC 283, and columns (5) to (8) report estimates for affiliates whose parent is in SIC 371. In Panel B, columns (1) to (4) report estimates for affiliates whose parent is in SIC 35, and columns (5) to (8) report estimates for affiliates whose parent is in SIC 28. Persistence corresponds to estimates of ρ . Affiliate R&D, Parent R&D, Other Affiliates' R&D and Affiliate R&D × Parent R&D estimates capture the elasticity of period t performance with respect to the period $t - 1$ value of the corresponding covariate. Labor Elasticity is the average value of $\beta_l + \beta_{ll}2l_{ijt} + \beta_{lk}k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk}2k_{ijt} + \beta_{lk}l_{ijt}$. The standard deviation for each of these input elasticities appears in parentheses below its mean. Measures of labor, capital, value added, and R&D expenditure are from the Bureau of Economic Analysis Survey of U.S. Direct Investment Abroad.

Table 9: Estimates with Affiliate Imports from U.S. Parent

	(1)	(2)	(3)	(4)
Persistence	0.7601 ^a (0.0099)	0.7568 ^a (0.0099)	0.7532 ^a (0.0099)	0.7520 ^a (0.0099)
Affiliate R&D	0.0029 ^b (0.0013)	-0.0248 ^a (0.0070)	0.0026 ^b (0.0013)	-0.0146 ^c (0.0077)
Parent R&D	0.0122 ^a (0.0023)	0.0070 ^a (0.0026)	0.0020 (0.0030)	0.0009 (0.0033)
Affiliate R&D \times Parent R&D		0.0022 ^a (0.0005)		0.0013 ^c (0.0007)
Parent Imports	0.0085 ^a (0.0012)	0.0082 ^a (0.0012)	-0.0185 ^a (0.0055)	-0.0140 ^a (0.0057)
Parent R&D \times Parent Imports			0.0022 ^a (0.0004)	0.0018 ^a (0.0005)
Affiliate R&D \times Parent R&D \times Parent Imports				0.0000 (0.0000)
Labor Elasticity	0.4682 (0.0646)	0.4682 (0.0646)	0.4682 (0.0646)	0.4682 (0.0646)
Capital Elasticity	0.2393 (0.0513)	0.2360 (0.0498)	0.2300 (0.0482)	0.2290 (0.0475)
Observations	4,168	4,168	4,168	4,168

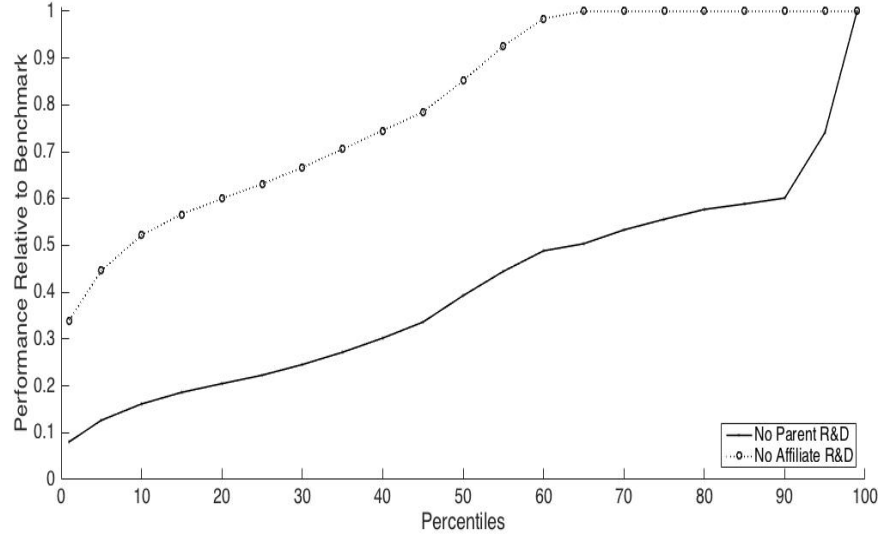
Notes: *a* denotes 1% significance, *b* denotes 5% significance, *c* denotes 10% significance. This table reports Nonlinear Least Squares estimates of (22) and several variants, incorporating observed affiliate imports from the U.S. parent (Parent Imports). All columns include market-year fixed effects. Standard errors are reported in parenthesis. Persistence corresponds to estimates of ρ . Labor Elasticity is the average value of $\beta_l + \beta_{ll}2l_{ijt} + \beta_{lk}k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk}2k_{ijt} + \beta_{lk}l_{ijt}$. The standard deviation for each of these input elasticities appears in parentheses below its mean. All other estimates capture the elasticity of period t performance with respect to the $t - 1$ value of the corresponding covariate. Measures of labor, capital, value added, R&D expenditure, and affiliate imports from the parent are from the Bureau of Economic Analysis Survey of U.S. Direct Investment Abroad.

Table 10: Misreporting of R&D Expenditures and Value-Added

	<i>No Tax Havens</i>			<i>Low Imports from Parent</i>		GMM		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Persistence	0.7648 ^a (0.0106)	0.7639 ^a (0.0106)	0.8036 ^a (0.0122)	0.8002 ^a (0.0122)	0.6749 ^a (0.0327)	0.7221 ^a (0.0354)	0.7109 ^a (0.0267)	0.6763 ^a (0.0338)
Affiliate R&D	0.0056 ^a (0.0014)	-0.0163 ^b (0.0076)	0.0046 ^b (0.0018)	-0.0247 ^a (0.0087)	0.0365 ^a (0.0122)	-0.2704 ^b (0.1140)	0.0293 ^a (0.0110)	0.2493 ^b (0.1215)
Parent R&D	0.0182 ^a (0.0024)	0.0139 ^a (0.0028)	0.0128 ^a (0.0027)	0.0069 ^a (0.0032)	0.0326 ^b (0.0161)	-0.0791 ^b (0.0377)	0.0322 ^c (0.0169)	0.0890 ^a (0.0345)
Other Affiliates' R&D								
Affiliate R&D × Parent R&D		0.0017 ^a (0.0006)		0.0024 ^a (0.0007)		0.0240 ^a (0.0087)		-0.0160 (0.0086)
Labor Elasticity	0.4813 (0.0678)	0.4813 (0.0678)	0.4525 (0.0722)	0.4525 (0.0722)	0.4682 (0.0646)	0.4682 (0.0646)	0.4682 (0.0646)	0.4682 (0.0646)
Capital Elasticity	0.2324 (0.0537)	0.2284 (0.0514)	0.1956 (0.0361)	0.1900 (0.0337)	0.2245 (0.0545)	0.2094 (0.0437)	0.2115 (0.0436)	0.2041 (0.0463)
Observations	3,432	3,432	2,212	2,212	4,194	4,194	4,194	4,194

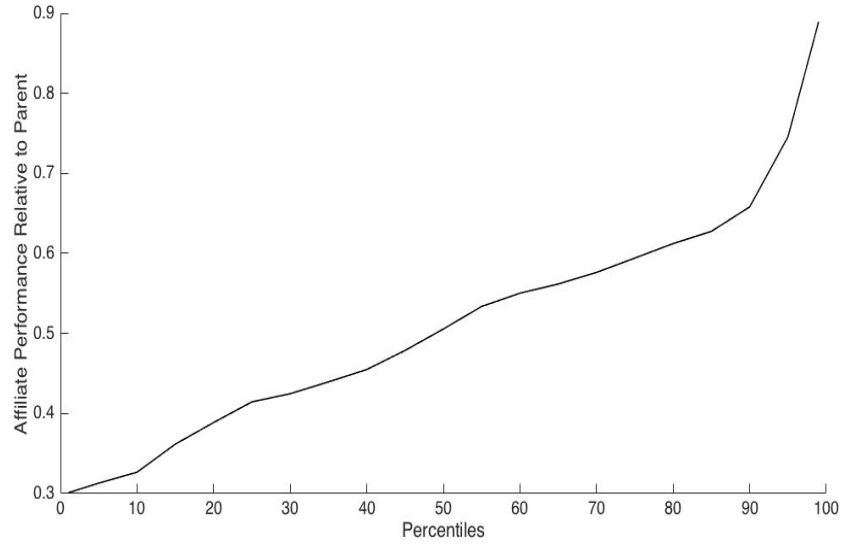
Notes: *a* denotes 1% significance, *b* denotes 5% significance, *c* denotes 10% significance. All columns report optimal two-step Generalized Method of Moments estimators of the same parameters. All columns include market-year fixed effects; columns (7) and (8) also include U.S. state fixed effects corresponding to the parent headquarters location. Standard errors are reported in parentheses. Persistence corresponds to estimates of ρ . Affiliate R&D, Parent R&D, Other Affiliates' R&D and Affiliate R&D \times Parent R&D estimates capture the elasticity of period t performance with respect to the period $t - 1$ value of the corresponding covariate. Labor Elasticity is the average value of $\beta_l + \beta_l 2l_{ijt} + \beta_{lk} k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_k 2k_{ijt} + \beta_{ll} l_{ijt}$. The standard deviation for each of these input elasticities appears in parentheses below its mean. The instrument for parent R&D is the user cost of R&D from Wilson (2009); the instrument for affiliate R&D is the interaction between a) the user cost of R&D prevailing in the U.S. headquarters state, and b) the strength of intellectual property rights in the affiliate country from Ginarte and Park (1997) and Park (2008). Measures of labor, capital, value added, and R&D expenditure are from the Bureau of Economic Analysis Survey of U.S. Direct Investment Abroad.

Figure 1: Determinants of Long-Run Affiliate Performance Distribution



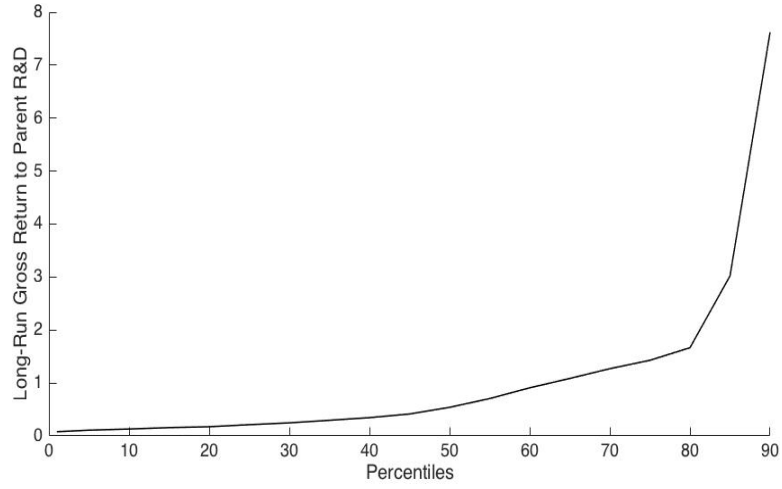
Notes: For each percentile indicated on the horizontal axis, the height of the solid line is the long-run affiliate performance without parent R&D relative to the long-run affiliate performance level in our benchmark specification; and the height of the dotted line with circles indicates long-run affiliate performance without affiliate R&D relative to the benchmark specification.

Figure 2: Distribution of Affiliate Performance Relative to Firm Parent



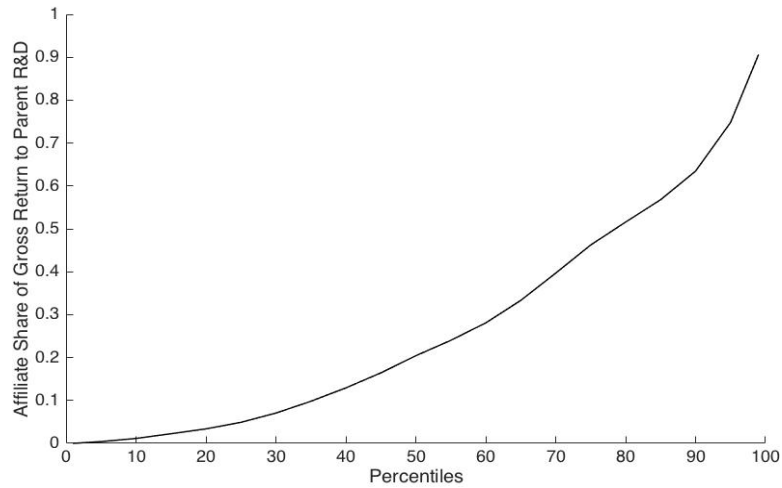
Notes: For each percentile indicated on the horizontal axis, the height of the solid line is the expected long-run affiliate performance due to affiliate and parent R&D spending and affiliate imports from the parent relative to the expected long-run parent performance due to parent R&D spending.

Figure 3: Distribution of Gross Return to Parent R&D



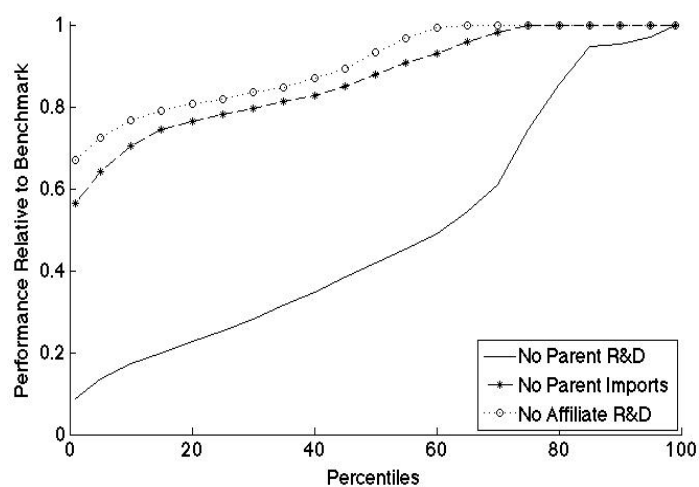
Notes: For each percentile indicated on the horizontal axis, the height of the solid line is the long-run gross return to the investment in R&D performed by U.S. parents in the year 2004.

Figure 4: Affiliate Share of Gross Return to Parent R&D



Notes: For each percentile indicated on the horizontal axis, the height of the solid line is the share of the long-run gross return to investment in R&D performed by U.S. parents in the year 2004 that can be attributed to affiliates.

Figure 5: Determinants of Long-Run Affiliate Performance Distribution Accounting for Intrafirm Trade



Notes: For each percentile indicated on the horizontal axis, the height of the solid line is the long-run affiliate performance without parent R&D relative to the long-run affiliate performance level in our benchmark specification; the height of the dashed line marked with asterisks indicates long-run affiliate performance without parent imports relative to the benchmark specification; and the height of the dotted line with circles indicates long-run affiliate performance without affiliate R&D relative to the benchmark specification.

Appendix (For Online Publication)

A.1 Value Added Function

This section includes a detailed derivation of (5). Assume firm i determines the optimal quantity of material inputs used for production by affiliate j in period t by maximizing the static profits in (12) with respect to M_{ijt} . Combining the definition of value added (revenue less expenditure on materials) with the demand function in (1), this maximization problem may be expressed as follows

$$\begin{aligned}\max_{M_{ijt}}\{Y_{ijt} - P_{n_{ijt}}^m M_{ijt}\} &= \max_{M_{ijt}}\{P_{ijt} Q_{ijt} - P_{n_{ijt}}^m M_{ijt}\} \\ &= \max_{M_{ijt}}\{Q_{n_{ijt}}^{\frac{1}{\sigma}} P_{n_{ijt}} \exp\left(\xi_{ijt} \frac{\sigma-1}{\sigma}\right) Q_{ijt}^{\frac{\sigma-1}{\sigma}} - P_{n_{ijt}}^m M_{ijt}\}.\end{aligned}$$

Given the production function in (2), the optimal level of material inputs M_{ijt}^* satisfies the following condition

$$P_{n_{ijt}}^m = \frac{\alpha_m(\sigma-1)}{\sigma} Q_{n_{ijt}}^{\frac{1}{\sigma}} P_{n_{ijt}} \exp\left[\frac{(\xi_{ijt} + \omega_{ijt})(\sigma-1)}{\sigma}\right] [H(K_{ijt}, L_{ijt}; \boldsymbol{\alpha})]^{\frac{(1-\alpha_m)(\sigma-1)}{\sigma}} M_{ijt}^{\frac{\alpha_m(\sigma-1)}{\sigma}-1}.$$

Thus, assuming firms determine materials use optimally, the revenue function may be rewritten in logs as

$$\begin{aligned}y_{ijt} &= \frac{\alpha_m(\sigma-1)}{\sigma - \alpha_m(\sigma-1)} \ln\left(\frac{\alpha_m(\sigma-1)}{\sigma}\right) - \frac{\alpha_m(\sigma-1)}{\sigma - \alpha_m(\sigma-1)} p_{n_{ijt}}^m \\ &\quad + \frac{\sigma}{\sigma - \alpha_m(\sigma-1)} p_{n_{ijt}} + \frac{1}{\sigma - \alpha_m(\sigma-1)} q_{n_{ijt}} \\ &\quad + \frac{(1-\alpha_m)(\sigma-1)}{\sigma - \alpha_m(\sigma-1)} h(k_{ijt}, l_{ijt}; \boldsymbol{\alpha}) + \frac{(\sigma-1)}{\sigma - \alpha_m(\sigma-1)} (\omega_{ijt} + \xi_{ijt})\end{aligned}$$

or, more concisely, as $y_{ijt} = \tilde{\kappa}_{n_{ijt}} + \iota(1-\alpha_m)h(k_{ijt}, l_{ijt}; \boldsymbol{\alpha}) + \psi_{ijt}$, where

$$\begin{aligned}\iota &= \frac{(\sigma-1)}{\sigma - \alpha_m(\sigma-1)}, \\ \tilde{\kappa}_{n_{ijt}} &= \iota \left[\alpha_m \ln\left(\frac{\alpha_m(\sigma-1)}{\sigma}\right) - \alpha_m p_{n_{ijt}}^m + \frac{\sigma}{\sigma-1} p_{n_{ijt}} + \frac{1}{\sigma-1} q_{n_{ijt}} \right], \\ \psi_{ijt} &= \iota(\omega_{ijt} + \xi_{ijt}).\end{aligned}$$

Because the translog function $h(k_{ijt}, l_{ijt}; \boldsymbol{\alpha})$ is linear in $\boldsymbol{\alpha}$, we can also represent the revenue function as

$$y_{ijt} = \tilde{\kappa}_{n_{ijt}} + h(k_{ijt}, l_{ijt}; \boldsymbol{\beta}) + \psi_{ijt},$$

where $\boldsymbol{\beta} = \boldsymbol{\alpha}\iota(1-\alpha_m)$.

Obtaining a similar expression for the value added function is straightforward. The first order condition for materials is

$$P_{n_{ijt}}^m M_{ijt}^* = \left(\frac{\alpha_m(\sigma-1)}{\sigma}\right) Y_{ijt}, \tag{A.1}$$

which implies that, conditional on M_{ijt}^* , value added is related to revenue as follows

$$VA_{ijt}^* = \left(1 - \frac{\alpha_m(\sigma-1)}{\sigma}\right) Y_{ijt}.$$

Value added (in logs) may thus be concisely represented as in (5),

$$va_{ijt}^* = \kappa_{n_{ijt}} + h(k_{ijt}, l_{ijt}; \beta) + \psi_{ijt},$$

where

$$\kappa_{n_{ijt}} = \ln \left(1 - \frac{\alpha_m(\sigma - 1)}{\sigma} \right) + \tilde{\kappa}_{n_{ijt}}.$$

A.2 Evolution of Performance: Details

Equations (8) and (9) describe the evolution over time of the performance index ψ_{ijt} :

$$\psi_{ijt} = \rho\psi_{ijt-1} + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \mu_{n_{ijt}} + \eta_{ijt}, \quad (\text{A.2})$$

As indicated in Section 2.4, this performance index is proportional to the sum of a productivity term ω_{ijt} and a demand shifter ξ_{ijt} : $\psi_{ijt} \equiv \iota(\omega_{ijt} + \xi_{ijt})$. Here we discuss two alternative specifications of stochastic processes for ω_{ijt} and ξ_{ijt} such that their sum scaled by ι is equal to the autoregressive process of order one in equation (A.2).

First, the stochastic process in equation (A.2) may be rationalized by both ω_{ijt} and ξ_{ijt} following autoregressive processes of order one with identical persistence parameter equal to ρ :

$$\omega_{ijt} = \rho\omega_{ijt-1} + \mu_a^\omega r_{ijt-1} + \mu_p^\omega r_{i0t-1} + \mu_{ap}^\omega r_{ijt-1} r_{i0t-1} + \mu_{n_{ijt}}^\omega + \eta_{ijt}^\omega, \quad \mathbb{E}_{t-1}[\eta_{ijt}^\omega] = 0, \quad (\text{A.3})$$

$$\xi_{ijt} = \rho\xi_{ijt-1} + \mu_a^\xi r_{ijt-1} + \mu_p^\xi r_{i0t-1} + \mu_{ap}^\xi r_{ijt-1} r_{i0t-1} + \mu_{n_{ijt}}^\xi + \eta_{ijt}^\xi, \quad \mathbb{E}_{t-1}[\eta_{ijt}^\xi] = 0. \quad (\text{A.4})$$

Summing up these two equations and multiplying them by ι , we obtain:

$$\begin{aligned} \psi_{ijt} = & \rho\psi_{ijt-1} + \iota(\mu_a^\omega + \mu_a^\xi) r_{ijt-1} + \iota(\mu_p^\omega + \mu_p^\xi) r_{i0t-1} + \iota(\mu_{ap}^\omega + \mu_{ap}^\xi) r_{ijt-1} r_{i0t-1} + \iota(\mu_{n_{ijt}}^\omega + \mu_{n_{ijt}}^\xi) \\ & + \iota(\eta_{ijt}^\omega + \eta_{ijt}^\xi), \end{aligned}$$

which is identical to equation (A.2) as long as

$$\mu_a \equiv \iota(\mu_a^\omega + \mu_a^\xi) \quad \mu_p \equiv \iota(\mu_p^\omega + \mu_p^\xi) \quad \mu_{ap} \equiv \iota(\mu_{ap}^\omega + \mu_{ap}^\xi) \quad \mu_{n_{ijt}} \equiv \iota(\mu_{n_{ijt}}^\omega + \mu_{n_{ijt}}^\xi) \quad \eta_{ijt} \equiv \iota(\eta_{ijt}^\omega + \eta_{ijt}^\xi).$$

It is important to note that the key restriction is that the autoregressive parameter is equal in equations (A.3) and (A.4), the reason being that the sum of two AR(1) processes with different persistence parameters is no longer an AR(1) but an ARMA(2,1). No restriction is imposed on the remaining parameters in these two equations. For example, it could be the case that parent R&D does not affect affiliates' demand level, $\mu_p^\xi = 0$, and that, therefore, the estimated impact of parent R&D on affiliate performance is all entirely due to the impact that it has on affiliates' productivity, $\mu_p = \iota\mu_p^\omega$.

Second, the stochastic process in equation (A.2) may be rationalized by either ω_{ijt} or ξ_{ijt} following an autoregressive process of order one with persistence parameter ρ and the other one having persistence equal to zero. For example:

$$\omega_{ijt} = \rho\omega_{ijt-1} + \mu_a^\omega r_{ijt-1} + \mu_p^\omega r_{i0t-1} + \mu_{ap}^\omega r_{ijt-1} r_{i0t-1} + \mu_{n_{ijt}}^\omega + \eta_{ijt}^\omega, \quad \mathbb{E}_{t-1}[\eta_{ijt}^\omega] = 0, \quad (\text{A.5})$$

$$\xi_{ijt} = \mu_a^\xi r_{ijt-1} + \mu_p^\xi r_{i0t-1} + \mu_{ap}^\xi r_{ijt-1} r_{i0t-1} + \mu_{n_{ijt}}^\xi + \eta_{ijt}^\xi, \quad \mathbb{E}_{t-1}[\eta_{ijt}^\xi] = 0. \quad (\text{A.6})$$

A.3 Data and Measurement

Multinational activity: Confidential data on U.S. multinational firms and their activity abroad is provided by the Bureau of Economic Analysis (BEA) through a sworn-status research arrangement. The data include detailed financial and operating information for each foreign affiliate owned (at least a 10% share) by a U.S. entity. In our estimation, we use information on the value added, labor (number of employees), capital (the value of plant, property, and equipment, net of depreciation), research and development (R&D) spending, R&D labor (number of R&D employees), and employee compensation corresponding, separately, to the U.S. parent and each of its affiliates abroad. These variables were extracted from the BEA’s comprehensive data files for each year, and then merged by parent and affiliate identification numbers to form a complete panel. The dataset used in the estimation covers U.S. affiliates during 1989–2008.

The measure of parent and affiliate-level value added used in our analysis is constructed by the BEA. This measure follows the definition in Mataloni and Goldberg (1994) from the factor-cost side, in which value added is employee compensation (wages and salaries plus employee benefits), plus profit-type returns (net income plus income taxes plus depreciation, less capital gains and losses, less income from equity investments), plus net interest paid (monetary interest paid plus imputed interest paid, less monetary interest received, less imputed interest received), plus indirect business taxes (taxes other than income and payroll taxes plus production royalty payments to governments, less subsidies received), plus capital consumption allowances (depreciation).

We measure affiliate-level capital as the value of plant, property, and equipment, net of depreciation corresponding to the affiliate site. This value is reported directly to the BEA in benchmark years and in each year immediately preceding a benchmark year. We therefore observe K_{ijt} directly in 1989, 1993, 1994, 1998, 1999, 2003, and 2004. For all remaining years, we construct affiliate-level capital by combining K_{ijt} values with observed investment in physical capital (plant, property, and equipment) using the perpetual inventory method and a depreciation rate of 5.9 percent, the physical capital depreciation rate found for U.S. manufacturing firms in Nadiri and Prucha (1996).

Research and development expenditures are reported directly and include basic and applied research in science and engineering, and the design and development of prototypes and processes, if the purpose of such activity is to: 1) pursue a planned search for new knowledge whether or not the search has reference to a specific application; 2) apply existing knowledge to the creation of a new product or process, including evaluation of use; or 3) apply existing knowledge to the employment of a present product or process. This variable includes all costs incurred to support R&D, including R&D depreciation and overhead. The variable excludes capital expenditures, routine product testing and quality control conducted during commercial production, geological and geophysical exploration, market research and surveys, and legal patent work.

All estimates in sections 5 and 6 arise from specifications that control for country-year fixed effects, but we nevertheless convert all variables originally expressed U.S. dollar nominal values to 2004 real terms using correction factors available from the U.S. Bureau of Labor Statistics.

We estimate model parameters separately by industry. The baseline estimates presented in section 5 correspond to multinational firms in the computer and office equipment industry (SIC 357). We define the industry of each multinational corporation based on the 3-digit SIC sector of its U.S. parent; i.e. for a parent that reports sales in a given 3-digit SIC sector, we extract all available observations for it and each of its manufacturing affiliates abroad during 1989–2008. Non-manufacturing affiliates, including those primarily operating in retail, finance, insurance, and agriculture are thus excluded. We also report estimates using a separate panel that employs a more restrictive industry definition whereby affiliates are included only if they, too, report sales in the same 3-digit SIC sector as its parent. Results for the computer and office equipment industry (SIC 357) based on this narrower definition of industry are described in section

6. Section 6 also reports results for two other three-digit SIC industries and the corresponding broader two-digit SIC industries: motor vehicles and motor vehicle equipment (SIC 371), pharmaceutical drugs (SIC 283), industrial and commercial machinery (SIC 35), transportation equipment (SIC 37), and chemicals (SIC 28), which are the three largest manufacturing industries based on the number of U.S. firm affiliates abroad (U.S. BEA 2008),⁶⁹

The data are cleaned prior to estimation. Observations are excluded if a) values are carried over or imputed based on previous survey responses; b) the affiliate is exempt from reporting R&D expenditures. Regarding b), the BEA requires only majority-owned and relatively large foreign affiliates of U.S. parent firms to report R&D expenditures. The reporting threshold differs depending on the year, ranging between \$3 million in 1989 and 1994 to \$50 million in 1999; thresholds in nominal terms were \$15 million in the period 1990–1993, \$25 million in 2004, \$20 million in 1995–1998, \$30 million in 2000–2003, and \$40 million in 2005–2008. We impose year-specific cutoffs to build the dataset used in the baseline estimation.

Table A.1: Descriptive Statistics, R&D Allocation in the Multinational Firm: Additional Industries

	SIC 28	SIC 35	SIC 37
Number of Affiliates per Firm with Positive R&D, Mean	5.07	2.11	2.85
Number of Affiliates per Firm, Mean	15.85	8.62	11.29
Affiliate Share in Total Firm R&D Expenditure, Aggregate	0.17	0.11	0.10
Affiliate Share in Total Firm Sales, Aggregate	0.42	0.41	0.36
Affiliate Share in Total Firm Employment, Aggregate	0.40	0.35	0.35

Notes: All variables are from the 1994 Bureau of Economic Analysis Benchmark Survey of U.S. Direct Investment Abroad; the 1994 survey is unusually comprehensive in its coverage of U.S. multinational firms' activity abroad. These data are a cross section covering U.S.-based firms operating in Chemicals (SIC 28), Industrial Machinery (SIC 35), and Transportation Equipment (SIC 37).

R&D user costs: The user cost of R&D investment is from Wilson (2009). Wilson (2009) constructs the state-specific user cost of R&D by extending the neoclassical Hall-Jorgenson formula for the cost of capital (Hall and Jorgenson 1967) to incorporate state and federal corporate taxes and R&D subsidies. Wilson (2009) provides this index for all 50 U.S. states and the District of Columbia annually during 1981–2006. The exact formula for state- i user cost of R&D is

$$\frac{1 - s(k_{it} + k_{ft}) - z(\tau_{it} + \tau_{ft})}{1 - (\tau_{it} + \tau_{ft})} [r_t + \delta],$$

where t indexes time, f indexes federal-level variables, r is the real interest rate, δ is the R&D capital depreciation rate, τ is the effective corporate income tax rate, k is the effective R&D tax credit rate, s is the share of R&D expenditures that qualify for R&D subsidies, and z is the present discounted value of tax depreciation allowances. Further details regarding the index and its construction are provided in Wilson (2009).

Patent rights index: The index of patent protection is published in Ginarte and Park (1997) and Park (2008). The index is available for 122 countries between 1960 and 2005 at five-year intervals. It is constructed as the sum of five sub-indexes corresponding to 1) enforcement, 2) coverage, 3) provisions for the loss of protection, 4) duration, and 5) membership in international intellectual property treaties. Each sub-index ranges between zero and one. Further details are provided in the two aforementioned publications.

⁶⁹During the sample period, the BEA data switches from using SIC to using NAICS-based parent-firm and foreign-affiliate industry classifications. We apply the U.S. Census Bureau concordance to match NAICS-based observations to each of the five SIC industries for which we estimate model parameters.

Tax Havens: The identification of countries as tax havens is from Gravelle (2015). This list was prepared by the U.S. Congressional Research Service and is similar to lists prepared by the Organization for Economic Cooperation and Development (OECD) and the U.S. Government Accountability Office (GAO). The list of countries classified as tax havens is: Andorra, Bahamas, Bahrain, Barbados, Bermuda, Costa Rica, Cyprus, the Dominican Republic, Pacific Ocean French Islands, Indian Ocean French Islands, Gibraltar, Hong Kong, Ireland, Jordan, Lebanon, Liechtenstein, Luxembourg, Macau, Maldives, Malta, Mauritius, Monaco, the Netherlands, the Netherlands Antilles, Panama, the Seychelles, Switzerland, Singapore, and British Overseas Territories.

A.4 Sample Selection Bias

Here, we discuss the impact of instantaneous entry and exit for the estimates resulting from the procedure described in section 4.1. As described in section 4, a necessary condition for consistency of the estimation procedure is that the data-generating process verifies the mean independence restriction in (17). In order to study the effect that the assumed instantaneous entry and exit have on the validity of this mean-independence condition, we proceed by first rewriting its conditioning set in terms of an equivalent set of covariates. From (6) and the definition of \widehat{va}_{ijt-1} as

$$\widehat{va}_{ijt-1} \equiv va_{ijt-1} - \hat{\beta}_l l_{ijt-1} - \hat{\beta}_{ll} l_{ijt-1}^2 - \hat{\beta}_{lk} l_{ijt-1} k_{ijt-1} - \hat{\varepsilon}_{ijt-1},$$

we can write

$$\psi_{ijt-1} + \kappa_{n_{ijt}} = \widehat{va}_{ijt-1} - \beta_k k_{ijt-1} - \beta_{kk} k_{ijt-1}^2.$$

Plugging this equality into (16) we can thus rewrite the conditional expectation in (17) as

$$\mathbb{E}[\eta_{ijt} | k_{ijt}, \psi_{ijt-1}, r_{ijt-1}, r_{i0t-1}, \mu_{n_{ijt}}, \kappa_{n_{ijt}}, j \in \mathcal{J}_{it-1}, j \in \mathcal{J}_{it}] = 0. \quad (\text{A.7})$$

Section 4.2 shows that, whether or not we assume that entry decisions are instantaneous, (A.7) implies

$$\mathbb{E}[\eta_{ijt} | k_{ijt}, \psi_{ijt-1}, r_{ijt-1}, r_{i0t-1}, \mu_{n_{ijt}}, \kappa_{n_{ijt}}, j \in \mathcal{J}_{it}] = 0. \quad (\text{A.8})$$

If firm- i 's decision about affiliate j exit at period t is instantaneous, then the variable $\mathbb{1}\{j \in \mathcal{J}_{it}\}$ in the conditioning set in (A.7) becomes a function of \mathbf{S}_{it} and, from (10), implicitly a function of k_{ijt} , $\kappa_{n_{ijt}}$, and ψ_{ijt} . Furthermore, from (8) and (9), we can rewrite ψ_{ijt} as a function of ψ_{ijt-1} , r_{ijt-1} , r_{i0t-1} , $\mu_{n_{ijt}}$ and η_{ijt} . Therefore, in sum, the variable $\mathbb{1}\{j \in \mathcal{J}_{it}\}$ becomes a function of all the elements in the conditioning set in (A.7).

According to the firm optimization problem described in section 2, one may conjecture that the optimal solution for the decision by firm i of having affiliate j incorporated at period t is characterized by a threshold rule: there is a critical productivity level $\bar{\psi}_{ijt}$ such that, if $\psi_{ijt} \leq \bar{\psi}_{ijt}$, affiliate j is not integrated in multinational i at period t , and the opposite is true if $\psi_{ijt} \geq \bar{\psi}_{ijt}$. The threshold value $\bar{\psi}_{ijt}$ will be a function of all the elements of the state vector \mathbf{S}_{it} other than ψ_{ijt} . According to this conjecture, one may rewrite (A.8) as

$$\mathbb{E}[\eta_{ijt} | k_{ijt}, \psi_{ijt-1}, r_{ijt-1}, r_{i0t-1}, \mu_{n_{ijt}}, \kappa_{n_{ijt}}, \psi_{ijt} \geq \bar{\psi}_{ijt}] = 0. \quad (\text{A.9})$$

By (8), η_{ijt} is mean independent of all the elements in \mathbf{S}_{it-1} . Therefore, conditional on the selection rule $\psi_{ijt} \geq \bar{\psi}_{ijt}$, η_{ijt} is independent of ψ_{ijt-1} , r_{ijt-1} , r_{i0t-1} , $\gamma_{n_{ijt}}$. Also, k_{ijt} is exclusively a function of the

state vector in period $t - 1$ and before. Therefore, conditional on the selection rule $\psi_{ijt} \geq \bar{\psi}_{ijt}$, η_{ijt} is also independent of k_{ijt} . Therefore, we can simplify (A.9) as

$$\mathbb{E}[\eta_{ijt} | \psi_{ijt} \geq \bar{\psi}_{ijt}] = 0. \quad (\text{A.10})$$

From (8) and (9), we can further rewrite this expression as

$$\mathbb{E}[\eta_{ijt} | \rho\psi_{ijt-1} + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \mu_{n_{ijt}} + \eta_{ijt} \geq \bar{\psi}_{ijt}] = 0. \quad (\text{A.11})$$

Therefore, as long as ρ , μ_a , μ_p and μ_{ap} are all positive, the higher ψ_{ijt-1} , r_{ijt-1} or r_{i0t-1} are, the lower η_{ijt} must be so that $\psi_{ijt} \geq \bar{\psi}_{ijt}$. This shows that, if the participation decision affecting the set of affiliates of firm i at period t were instantaneous and thus taken after the state vector \mathbf{S}_{it} is realized (instead of being taken immediately after \mathbf{S}_{it-1} is realized, as it is assumed in section 2.6), the estimates of ρ , μ_a , μ_p and μ_{ap} obtained through the estimation procedure described in section 4.1 will be biased downward and will thus underestimate both the persistence of firm performance as well as the impact of parent and affiliate R&D on performance.

A.5 Heterogeneous Markups

One may decompose the scaling factor ι entering the definition of β and firm performance ψ_{ijt} (see section 2.4) into a component that exclusively depends on α_m and a component that combines α_m and σ as follows

$$\iota = \iota_1 \iota_2, \quad \text{with} \quad \iota_1 = \frac{1}{\alpha_m} \quad \text{and} \quad \iota_2 = \frac{\alpha_m \frac{\sigma-1}{\sigma}}{1 - \alpha_m \frac{\sigma-1}{\sigma}}. \quad (\text{A.12})$$

We describe here a procedure to estimate the parameter vector of interest when the demand elasticity σ and, therefore, the term ι_2 , varies by both market n and year t . We denote the market-year varying parameter σ as $\sigma_{n_{ijt}}$, and similarly denote the market-year varying parameter ι_2 as $\iota_{2n_{ijt}}$.

The procedure to estimate the parameters necessary to determine the short- and long-run impact of affiliate and parent R&D investment on affiliates' performance has three steps: (1) estimating $\iota_{2n_{ijt}}$ for every market n and year t ; (2) rewriting both the value added function and the stochastic process determining the evolution of affiliates' performance as a function of data, the estimates $\{\hat{\iota}_{2n_{ijt}}\}$, and a vector of parameters that do not vary by market or year; (3) estimating this vector of parameters.

Step 1: Estimating $\iota_{2n_{ijt}}$. Rearranging terms in (A.1), we obtain

$$\alpha_m \frac{\sigma_{n_{ijt}} - 1}{\sigma_{n_{ijt}}} = \frac{P_{n_{ijt}}^m M_{ijt}}{Y_{ijt}^*} = \frac{Y_{ijt}^* - VA_{ijt}^*}{Y_{ijt}^*},$$

where VA_{ijt}^* is defined section 2.4 and Y_{ijt}^* denotes the actual revenue of affiliate j of firm i in period t . Taking into account the measurement error in value added, ε_{ijt} , and allowing also for an analogous multiplicative measurement error affecting revenue, so that $Y_{ijt} \equiv Y_{ijt}^* \exp(\varepsilon_{ijt}^y)$, we write

$$\frac{VA_{ijt} \exp(-\varepsilon_{ijt})}{Y_{ijt} \exp(-\varepsilon_{ijt}^y)} = 1 - \alpha_m \frac{\sigma_{n_{ijt}} - 1}{\sigma_{n_{ijt}}}.$$

Taking logs, this becomes

$$va_{ijt} - y_{ijt} + (\varepsilon_{ijt}^y - \varepsilon_{ijt}) = \ln \left(1 - \alpha_m \frac{\sigma_{n_{ijt}} - 1}{\sigma_{n_{ijt}}} \right)$$

or, equivalently,

$$va_{ijt} - y_{ijt} = \ln \left(1 - \alpha_m \frac{\sigma_{n_{ijt}} - 1}{\sigma_{n_{ijt}}} \right) + (\varepsilon_{ijt} - \varepsilon_{ijt}^y).$$

Given the expression for ι_2 in equation (A.12), this may be expressed as

$$va_{ijt} - y_{ijt} = \ln \left(\frac{1}{1 + \iota_{2n_{ijt}}} \right) + (\varepsilon_{ijt} - \varepsilon_{ijt}^y).$$

Using D_{nt} to denote a dummy variable that takes the value 1 for market n and year t (and is otherwise zero), and assuming that both error terms have expectation zero, conditional on a market and year,

$$\mathbb{E}[\varepsilon_{ijt} D_{n_{ijt}}] = \mathbb{E}[\varepsilon_{ijt}^y D_{n_{ijt}}] = 0,$$

we obtain a consistent estimator of $\iota_{2n_{ijt}}$ for each nt pair using NLS and the following moment condition

$$\mathbb{E} \left[\left(va_{ijt} - y_{ijt} - \ln \left(\frac{1}{1 + \iota_{2n_{ijt}}} \right) \right) D_{n_{ijt}} \right] = 0.$$

We denote this estimator $\hat{\iota}_{2n_{ijt}}$.

Step 2: Value added function and evolution of firm performance conditional on $\hat{\iota}_{2n_{ijt}}$. Making use of $\hat{\iota}_{2n_{ijt}}$ and following the same steps as in section 2.4 we can rewrite the value added function in (6) as

$$va_{ijt} = \kappa_{n_{ijt}} + \tilde{h}(l_{ijt}, k_{ijt}, \hat{\iota}_{2n_{ijt}}; \tilde{\beta}) + \tilde{\psi}_{ijt} + \varepsilon_{ijt} \quad (\text{A.13})$$

where

$$\tilde{h}(l_{ijt}, k_{ijt}, \hat{\iota}_{2n_{ijt}}; \tilde{\beta}) = \tilde{\beta}_l \hat{\iota}_{2n_{ijt}} l_{ijt} + \tilde{\beta}_l \hat{\iota}_{2n_{ijt}} l_{ijt}^2 + \tilde{\beta}_k \hat{\iota}_{2n_{ijt}} k_{ijt} + \tilde{\beta}_{kk} \hat{\iota}_{2n_{ijt}} k_{ijt}^2 + \tilde{\beta}_{lk} \hat{\iota}_{2n_{ijt}} l_{ijt} k_{ijt}, \quad (\text{A.14})$$

$\tilde{\beta} = \alpha(1 - \alpha_m)\iota_1$ and $\tilde{\psi}_{ijt} = \iota_1 \hat{\iota}_{2n_{ijt}}(\omega_{ijt} + \xi_{ijt})$. Similarly, following the same steps as in section 2.5, we can write the modified firm performance $\tilde{\psi}_{ijt}$ as

$$\tilde{\psi}_{ijt} = \mathbb{E}_{t-1}[\tilde{\psi}_{ijt}] + \tilde{\eta}_{ijt}, \quad (\text{A.15})$$

where the expected period- t performance of affiliate j , conditional on the state vector of its multinational firm i at $t - 1$, is

$$\mathbb{E}_{t-1}[\tilde{\psi}_{ijt}] = \tilde{\rho} \hat{\iota}_{2n_{ijt}} \psi_{ijt-1} + \tilde{\mu}_a \hat{\iota}_{2n_{ijt}} r_{ijt-1} + \tilde{\mu}_p \hat{\iota}_{2n_{ijt}} r_{i0t-1} + \tilde{\mu}_{ap} \hat{\iota}_{2n_{ijt}} r_{ijt-1} r_{i0t-1} + \tilde{\mu}_{n_{ijt}}. \quad (\text{A.16})$$

Step 3: Estimating $(\tilde{\beta}, \tilde{\rho}, \tilde{\mu}_a, \tilde{\mu}_p, \tilde{\mu}_{ap}, \{\tilde{\gamma}_{n_{ijt}}\})$. Combining equations (A.13), (A.15), and (A.16), we obtain

$$\begin{aligned} va_{ijt} &= \tilde{h}(l_{ijt}, k_{ijt}, \hat{\iota}_{2n_{ijt}}; \tilde{\beta}) + \tilde{\rho}(va_{ijt-1} - \tilde{h}(l_{ijt-1}, k_{ijt-1}, \hat{\iota}_{2n_{ijt-1}}; \tilde{\beta})) \\ &\quad + \tilde{\mu}_a \hat{\iota}_{2n_{ijt}} r_{ijt-1} + \tilde{\mu}_p \hat{\iota}_{2n_{ijt}} r_{i0t-1} + \tilde{\mu}_{ap} \hat{\iota}_{2n_{ijt}} r_{ijt-1} r_{i0t-1} + \tilde{\gamma}_{n_{ijt}} + \tilde{u}_{ijt}, \end{aligned} \quad (\text{A.17})$$

where the error term is $u_{ijt} = \tilde{\eta}_{ijt} + \varepsilon_{ijt} - \tilde{\rho}\varepsilon_{ijt-1}$, and $\tilde{\gamma}_{n_{ijt}} = \tilde{\mu}_{n_{ijt}} + \kappa_{n_{ijt}} - \tilde{\rho}\kappa_{n_{ijt-1}}$. Given (A.17), we can follow analogous steps to those described in section 4.1 in order to obtain consistent estimates of $(\tilde{\beta}, \tilde{\rho}, \tilde{\mu}_a, \tilde{\mu}_p, \tilde{\mu}_{ap}, \{\tilde{\gamma}_{n_{ijt}}\})$.

Notice that we can use the estimates of the parameter vector $(\tilde{\beta}, \tilde{\rho}, \tilde{\mu}_a, \tilde{\mu}_p, \tilde{\mu}_{ap}, \{\tilde{\gamma}_{n_{ijt}}\})$, those of $\{\hat{\iota}_{2n_{ijt}}\}$

and (A.16) to compute the elasticity of $\tilde{\psi}_{ijt}$ with respect to r_{ijt-1} as

$$\hat{\mu}_a \hat{\ell}_{2n_{ijt}} + \hat{\mu}_{ap} \hat{\ell}_{2n_{ijt}} r_{i0t-1},$$

and the elasticity of $\tilde{\psi}_{ijt}$ with respect to r_{i0t-1} as

$$\hat{\mu}_p \hat{\ell}_{2n_{ijt}} + \hat{\mu}_{ap} \hat{\ell}_{2n_{ijt}} r_{ijt-1}.$$

In summary, the estimation approach described here allows us to use affiliate-level data on sales revenue and value added in a way that permits all parameters to vary by market and year. Specifically, it estimates first the component of these parameters that depends on the elasticity of demand and that, therefore, might vary across markets and years: $\{\hat{\ell}_{2n_{ijt}}\}$. Given these terms, it then follows a procedure very similar to that described in section 4.1 to estimate the components of the parameters determining the elasticity of output with respect to labor and capital and of performance with respect to lagged performance and R&D investment, which do not vary across market and years. The limited number of observations for each market-year pair in the data prevent estimating a more flexible model in which $(\beta, \mu_a, \mu_p, \mu_{ap})$ may vary across market and years in a way that is not constrained by the multiplicative terms $\{\hat{\ell}_{2n_{ijt}}\}$. We implement the procedures described here and report resulting estimates for the case in which σ may vary by year in section 6.

A.6 Variable Markups

Here we discuss alternative estimation approaches that differ from that in section 4.1 in that they yield consistent estimates of the parameter vector of interest without having to rely on the assumption of monopolistic competition. Specifically, we discuss three alternative estimation approaches. Given data availability, in all these three cases, relaxing the assumption of monopolistic competition will require restricting the production function defined in (2), (3), and (4). The first approach assumes away materials in (2) by assuming that $\alpha_m = 0$. The second approach additionally assumes that the value added function is Cobb-Douglas in labor and capital; i.e. $\alpha_{ll} = \alpha_{kk} = \alpha_{lk} = 0$. The third approach maintains the production function in (2), (3), and (4) but changes the definition of the variable M_{ijt} and interprets it as total expenditure on materials by affiliate j of firm i in period t .

Alternative Estimation Approach 1. Assume the demand function is defined by (1) and the production function is as described in (2) except for setting $\alpha_m = 0$. In this case, the revenue function is identical to the value added function and may be expressed in logs as

$$y_{ijt} = \kappa_{n_{ijt}} + h(l_{ijt}, k_{ijt}; \beta) + \psi_{ijt} + \varepsilon_{ijt}, \quad (\text{A.18})$$

where $\kappa_{n_{ijt}} = (1/\sigma)q_{n_{ijt}} + p_{n_{ijt}}$, $\beta = \alpha\iota$, $\psi_{ijt} = \iota(\omega_{ijt} + \xi_{ijt})$ and $\iota = (\sigma - 1)/\sigma$.⁷⁰ Combining this equation with the expressions for the evolution of firm performance in (8) and (9), we obtain an estimating equation that is analogous to that in (13):

$$\begin{aligned} y_{ijt} = h(k_{ijt}, l_{ijt}; \beta) + \rho(y_{ijt-1} - h(k_{ijt-1}, l_{ijt-1}; \beta)) + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} \\ + \mu_{ap} r_{ijt-1} r_{i0t-1} + \gamma_{n_{ijt}} + u_{ijt}, \end{aligned} \quad (\text{A.19})$$

with $u_{ijt} = \eta_{ijt} + \varepsilon_{ijt} - \rho\varepsilon_{ijt-1}$, and $\gamma_{n_{ijt}} = \mu_{n_{ijt}} + \kappa_{n_{ijt}} - \rho\kappa_{n_{ijt-1}}$.

⁷⁰In order to stress the similarities between the model in section 2 and estimation approach in 4.1 and those described here, we use the same symbols $\kappa_{n_{ijt}}$, β , ι and ψ_{ijt} even though they express slightly different variables.

Once we drop the assumption of monopolistic competition, equation (14) no longer captures the first order condition with respect to labor. In fact, this first order condition will depend on the equilibrium markup that affiliate j of firm i sets at period t . Following the cost minimization approach (see De Loecker and Warzynski 2012, and De Loecker et al 2016) to derive the first order condition for labor, we obtain:

$$\begin{aligned} \min_{L_{ijt}} \mathcal{L}(L_{ijt}) &= P_{ijt}^l L_{ijt} + P_{ijt}^k K_{ijt} + \lambda_{ijt}(Q_{ijt} - Q_{ijt}(\cdot)), \\ \frac{\partial \mathcal{L}_{it}}{\partial L_{ijt}} &= P_{ijt}^l - \lambda_{ijt} \frac{\partial Q_{ijt}(\cdot)}{\partial L_{ijt}} = 0, \end{aligned}$$

where λ_{ijt} is the Lagrange multiplier and all other variables are defined as in the main text. Rearranging terms and multiplying by L_{ijt}/Q_{ijt} , we obtain

$$\frac{\partial Q_{ijt}(\cdot)}{\partial L_{ijt}} \frac{L_{ijt}}{Q_{ijt}} = \frac{W_{ijt}^l}{Q_{ijt}} \frac{1}{\lambda_{ijt}}.$$

Multiplying and dividing by P_{ijt} in the right hand side, and noting that the Lagrange multiplier is equal to the marginal cost of production, we can write

$$\frac{\partial Q_{ijt}(\cdot)}{\partial L_{ijt}} \frac{L_{ijt}}{Q_{ijt}} = \frac{W_{ijt}^l}{Y_{ijt}^*} \zeta_{ijt},$$

where Y_{ijt}^* denotes the equilibrium revenue of affiliate j of firm i in period t , and ζ_{ijt} is the equilibrium markup, defined as the output price over the marginal cost. From the production function in (4), we can rewrite this first order condition with respect to labor as

$$\alpha_l + \alpha_{ll} 2l_{ijt} + \alpha_{lk} k_{ijt} = \frac{W_{ijt}^l}{Y_{ijt}^*} \zeta_{ijt},$$

and, multiplying by ι on both sides of this equality, we obtain

$$\beta_l + \beta_{ll} 2l_{ijt} + \beta_{lk} k_{ijt} = \frac{W_{ijt}^l}{Y_{ijt}^*} \tilde{\zeta}_{ijt},$$

where $\tilde{\zeta}_{ijt} = \iota \zeta_{ijt}$. Finally, allowing for measurement error in observed revenue as in the main text:

$$\beta_l + \beta_{ll} 2l_{ijt} + \beta_{lk} k_{ijt} = \frac{W_{ijt}^l}{Y_{ijt}} \exp(\varepsilon_{ijt}) \tilde{\zeta}_{ijt}. \quad (\text{A.20})$$

This expression is similar to that in (15) but there is a key difference. The reason why the moment condition in (15) is incompatible with the presence of variable markups is that it implies that the composite $\varepsilon_{ijt} + \ln(\tilde{\zeta}_{ijt})$ is mean independent of labor and capital. This will not be true in a general model of endogenous markups; the equilibrium markup of affiliate j in period t will be correlated with the quantity of inputs it hires.

Given that we cannot use equation (A.20) as an estimating for $(\beta_l, \beta_{ll}, \beta_{lk})$, we must estimate these parameters jointly with the remaining parameters using orthogonality conditions that apply to (A.19). We cannot use NLS to estimate the parameters entering (A.19) because both l_{ijt} and y_{ijt-1} will be correlated with the error term u_{ijt} : l_{ijt} is correlated with η_{ijt} and y_{ijt-1} is correlated with ε_{ijt-1} . Therefore, we need to find instruments for both l_{ijt} and y_{ijt-1} and use a GMM procedure to estimate the parameter vector $(\beta_l, \beta_{ll}, \beta_{lk}, \beta_k, \beta_{kk}, \rho, \mu_a, \mu_p, \mu_{ap})$.

Given that, from (7) and (8), y_{ijt-2} is mean independent of u_{ijt} , we may use this variable as instrument for y_{ijt-1} . This instrument is likely to be strong, as revenue tends to be persistent over time. There are two

variables in our data that may be used as instruments to identify the coefficients multiplying l_{ijt} in (A.17). From (7) and (8), l_{ijt-2} will be a valid instrument. However, once we condition on l_{ijt-1} , which enters directly in (A.17), the correlation between l_{ijt-2} and l_{ijt} is negligible in our data. An alternative would be to use w_{ijt} . However, for this instrument to be strong, there must be enough variation in wages paid by different affiliates located in the same market and year. At the same time, for w_{ijt} to be a valid instrument, we need that the wages paid by different affiliates are not correlated with their own productivity.⁷¹

Regardless of the moment conditions used to estimate the parameter vector of interest, note that the estimates of (μ_a, μ_p, μ_{ap}) would capture the effect that R&D expenditure has on firm performance. They do not capture the possible effect that R&D expenditure might have on markups. However, given estimates of β_l, β_{ll} and β_{lk} , we can measure the log of the composite of measurement error and markups for each affiliate as

$$\ln(\tilde{\zeta}_{ijt}) + \varepsilon_{ijt} = \ln(\beta_l + \beta_{ll}2l_{ijt} + \beta_{lk}k_{ijt}) - w_{ijt}^l + y_{ijt},$$

and we can study the impact that R&D expenditures has on markups by projecting $\ln(\tilde{\zeta}_{ijt}) + \varepsilon_{ijt}$ on different measures of R&D spending within a multinational firm i . Given that ε_{ijt} is assumed to be mean independent of any variable in the state vector \mathbf{S}_{it} (see (10)), the estimates of a regression of $\ln(\tilde{\zeta}_{ijt}) + \varepsilon_{ijt}$ on variables that are either included in \mathbf{S}_{it} or are a function of it will converge to the same values as the estimates of a regression of $\ln(\tilde{\zeta}_{ijt})$ on the same set of covariates. That is, ε_{ijt} would operate in such a regression as measurement error in the dependent variable and, therefore, would not affect the consistency of the coefficient estimates.

Summing up, this alternative estimation approach shows that we can relax the monopolistic competition assumption and thus allow for variable markups as long as, instead, we impose two additional assumptions not imposed in the model and estimation approach described in section 2 and 4.1: (a) production function does not depend on material usage; (b) wages paid by affiliates are mean independent of their performance indices. Assumption (a) is required because we do not observe investment in material inputs in our data. Assumption (b) is required if we use wages as instrument for labor usage in the GMM estimation of the parameter vector of interest or if we rely on the first order condition in (A.20) to recover affiliates' markups.

Alternative Estimation Approach 2. Assume the demand function is defined by (1) and the production function is as described in (2) except for setting $\alpha_m = \alpha_{ll} = \alpha_{lk} = \alpha_{kk} = 0$. In this case, the revenue function is identical to the value added function and may be expressed in logs as

$$y_{ijt} = \kappa_{n_{ijt}} + \beta_l l_{ijt} + \beta_k k_{ijt} + \psi_{ijt} + \varepsilon_{ijt}, \quad (\text{A.21})$$

where $\kappa_{n_{ijt}} = (1/\sigma)q_{n_{ijt}} + p_{n_{ijt}}$, $\beta_l = \alpha_{ll}$, $\beta_k = \alpha_{kk}$, $\psi_{ijt} = \iota(\omega_{ijt} + \xi_{ijt})$ and $\iota = (\sigma - 1)/\sigma$. Combining this equation with the expressions for the evolution of firm performance in (8) and (9), we obtain an estimating equation that is analogous to that in (13):

$$\begin{aligned} y_{ijt} = & \beta_l l_{ijt} + \beta_k k_{ijt} + \rho(y_{ijt-1} - \beta_l l_{ijt-1} - \beta_k k_{ijt-1}) + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} \\ & + \mu_{ap} r_{ijt-1} r_{i0t-1} + \gamma_{n_{ijt}} + u_{ijt}, \end{aligned} \quad (\text{A.22})$$

with $u_{ijt} = \eta_{ijt} + \varepsilon_{ijt} - \rho\varepsilon_{ijt-1}$, and $\gamma_{n_{ijt}} = \mu_{n_{ijt}} + \kappa_{n_{ijt}} - \rho\kappa_{n_{ijt-1}}$. Following the same steps as in Approach

⁷¹Notice that the coefficients entering (A.19) multiplying some term that depends on l_{ijt} -i.e. $(\beta_l, \beta_{ll}, \beta_{lk})$ - also enter multiplying terms that depend on l_{ijt-1} . Therefore, one could theoretically estimate the parameter vector of interest $(\beta_l, \beta_{ll}, \beta_{lk}, \beta_k, \beta_{kk}, \rho, \mu_a, \mu_p, \mu_{ap})$ without relying on moment conditions that use as instruments covariates that are both correlated with l_{ijt} and assumed to be mean independent of the error term u_{ijt} .

1 above, we obtain the following expression for the first order condition with respect to labor,

$$\beta_l = \frac{W_{ijt}^l}{Y_{ijt}} \exp(\varepsilon_{ijt}) \tilde{\zeta}_{ijt}. \quad (\text{A.23})$$

The advantage of assuming a Cobb-Douglas production function is that we can study the impact of R&D expenditures on markups without having to first compute consistent estimates of the parameters determining the elasticity of revenue with respect to labor; i.e. β_l . The reason is that, once we include a constant in the regression of $\ln(\tilde{\zeta}_{ijt}) + \varepsilon_{ijt} - \ln(\beta_l)$ on measures of R&D expenditures, the estimates of the coefficients on these measures of R&D spending will converge to the same values to which they would converge if the dependent variable were to be only $\ln(\tilde{\zeta}_{ijt})$. The reason for this is: (1) ε_{ijt} operates as measurement error in the dependent variable and, therefore, does not affect the consistency of estimates of regression coefficients; (2) $\ln(\beta_l)$ is just a constant and, therefore, the coefficients of a regression that has $\ln(\tilde{\zeta}_{ijt}) + \varepsilon_{ijt} - \ln(\beta_l)$ as dependent variable will differ from those of a regression that has $\ln(\tilde{\zeta}_{ijt}) + \varepsilon_{ijt}$ as dependent variable only in the constant term.

In order to estimate the parameter vector $(\beta_l, \beta_k, \rho, \mu_a, \mu_p, \mu_{ap})$, we use (A.22). As discussed above in Approach 1, obtaining a consistent estimate of $(\beta_l, \beta_k, \rho, \mu_a, \mu_p, \mu_{ap})$ using (A.22) as the estimating equation requires obtaining an instrument for y_{ijt-1} and potentially also for l_{ijt} . Also as discussed above, obtaining an instrument for l_{ijt} that is both valid and strong may be challenging.

In summary, assuming that the production function is Cobb-Douglas in labor and capital does not simplify the estimation procedure that one must follow to estimate both the production function parameters and the parameters determining the impact of R&D investment on performance.⁷² However, it significantly simplifies the estimation of the parameters determining the impact that R&D investment has on affiliates' markups; the reason being that these parameters may be estimated without having to previously estimate the parameters entering the production function, (β_l, β_k) , or the parameters determining the evolution of affiliate performance, $(\rho, \mu_a, \mu_p, \mu_{ap})$.

Alternative Estimation Approach 3. Assume the demand function is defined by (1) and the production function is as described in (2):

$$Q_{ijt} = (H(K_{ijt}, L_{ijt}; \boldsymbol{\alpha}))^{1-\alpha_m} (M_{ijt}^*)^{\alpha_m} \exp(\omega_{ijt}), \quad (\text{A.24})$$

where here we use M_{ijt}^* to denote the total expenditure in materials. In this case, by definition, $M_{ijt}^* = Y_{ijt}^* - VA_{ijt}^*$, where recall that Y_{ijt}^* and VA_{ijt}^* denote sales revenue and value added, respectively, for affiliate j of firm i in period t . Allowing for measurement error in both sales revenue, $Y_{ijt} = Y_{ijt}^* \exp(\varepsilon_{ijt}^y)$, and value added, $Y_{ijt} = Y_{ijt}^* \exp(\varepsilon_{ijt}^{va})$, we can write our measure of the log of materials use by affiliate j of firm i in period t as

$$m_{ijt} = y_{ijt} - va_{ijt} + \varepsilon_{ijt}^y - \varepsilon_{ijt}^{va}. \quad (\text{A.25})$$

In this case, the revenue function becomes

$$y_{ijt} = \kappa_{n_{ijt}} + h(l_{ijt}, k_{ijt}; \boldsymbol{\beta}) + \beta_m m_{ijt} + \psi_{ijt} + \varepsilon_{ijt}^y, \quad (\text{A.26})$$

where $\kappa_{n_{ijt}} = (1/\sigma)q_{n_{ijt}} + p_{n_{ijt}}$, $\boldsymbol{\beta} = \boldsymbol{\alpha}$, $\beta_m = \alpha_m \iota$, $\psi_{ijt} = \iota(\omega_{ijt} + \xi_{ijt})$ and $\iota = (\sigma - 1)/\sigma$. Combining

⁷²Even though the procedure is the same, the set of production function parameters to estimate is obviously smaller: (β_l, β_k) vs. $(\beta_l, \beta_k, \beta_{ll}, \beta_{lk}, \beta_{kk})$. In practice, in many dataset, it might be complicated to precisely estimate the additional parameters of the translog production function. Therefore, in empirical applications, assuming a Cobb-Douglas production function might significantly simplify estimating the parameters of interest.

(A.25) and (A.26), we can rewrite revenue as

$$\begin{aligned} y_{ijt} &= \kappa_{n_{ijt}} + h(l_{ijt}, k_{ijt}; \beta) + \beta_m(y_{ijt} - va_{ijt}) + \psi_{ijt} + (1 + \beta_m)\varepsilon_{ijt}^y - \beta_m\varepsilon_{ijt}^{va} \\ &= \tilde{\kappa}_{n_{ijt}} + h(l_{ijt}, k_{ijt}; \tilde{\beta}) - (\beta_m/(1 - \beta_m))va_{ijt} + \tilde{\psi}_{ijt} + ((1 + \beta_m)/(1 - \beta_m))\varepsilon_{ijt}^y - (\beta_m/(1 - \beta_m))\varepsilon_{ijt}^{va} \end{aligned} \quad (\text{A.27})$$

where $\tilde{\kappa}_{n_{ijt}} = (1/(1 - \beta_m))\kappa_{n_{ijt}}$, $\tilde{\beta} = (1/(1 - \beta_m))\beta$, $\tilde{\psi}_{ijt} = (1/(1 - \beta_m))\psi_{ijt}$. Combining this equation with the expressions for the evolution of firm performance in (8) and (9), we obtain:

$$\begin{aligned} y_{ijt} &= h(k_{ijt}, l_{ijt}; \tilde{\beta}) - (\beta_m/(1 - \beta_m))va_{ijt} + \rho(y_{ijt-1} - h(k_{ijt-1}, l_{ijt-1}; \tilde{\beta}) - (\beta_m/(1 - \beta_m))va_{ijt-1}) \\ &\quad + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \gamma_{n_{ijt}} + u_{ijt}, \end{aligned} \quad (\text{A.28})$$

with $u_{ijt} = \eta_{ijt} + ((1 + \beta_m)/(1 - \beta_m))\varepsilon_{ijt}^y - (\beta_m/(1 - \beta_m))\varepsilon_{ijt}^{va} - \rho(((1 + \beta_m)/(1 - \beta_m))\varepsilon_{ijt-1}^y - (\beta_m/(1 - \beta_m))\varepsilon_{ijt-1}^{va})$, and $\gamma_{n_{ijt}} = \mu_{n_{ijt}} + \tilde{\kappa}_{n_{ijt}} - \rho\tilde{\kappa}_{n_{ijt-1}}$. Following the same steps as in Approach 1 above, we obtain the following expression for the first order condition with respect to labor,

$$\beta_l l_{ijt} + \beta_{ll} 2l_{ijt} + \beta_{lk} k_{ijt} = \frac{W_{ijt}^l}{Y_{ijt}} \exp(\varepsilon_{ijt}^y) \tilde{\zeta}_{ijt}, \quad (\text{A.29})$$

where recall that ζ_{ijt} is the equilibrium markup and $\tilde{\zeta}_{ijt} = \iota \zeta_{ijt}$. The first order condition with respect to materials is

$$\beta_m = \frac{M_{ijt}}{Y_{ijt}} \exp(\varepsilon_{ijt}^y) \tilde{\zeta}_{ijt} = \frac{Y_{ijt} \exp(-\varepsilon_{ijt}^y) - V A_{ijt} \exp(-\varepsilon_{ijt}^{va})}{Y_{ijt} \exp(-\varepsilon_{ijt}^y)} \tilde{\zeta}_{ijt},$$

or, equivalently,

$$1 - \beta_m \frac{1}{\tilde{\zeta}_{ijt}} = \frac{V A_{ijt}}{Y_{ijt}} \exp(\varepsilon_{ijt}^y - \varepsilon_{ijt}^{va}). \quad (\text{A.30})$$

In this framework, we must estimate the parameter vector of interest $(\beta_l, \beta_{ll}, \beta_{lk}, \beta_k, \beta_{kk}, \beta_m, \rho, \mu_a, \mu_p, \mu_{ap})$ using (A.28) as estimating equation. As discussed already above in Approach 1, deriving orthogonality restrictions from (A.28) for estimation requires instruments for y_{ijt-1} and potentially also for l_{ijt} . In addition, identifying β_m using (A.28) requires instruments for va_{ijt-1} and potentially also for va_{ijt} ; finding two separate instruments, one for va_{ijt} and another one for va_{ijt-1} may be challenging.⁷³ Regardless of the instrumentation strategy, given estimates of β_l , β_{ll} and β_{lk} , it is possible to measure the log of composite of measurement error and markups for each affiliate from (A.29). As this equation is identical to (A.20), we refer here to the discussion in Approach 1 above on how to identify the impact of R&D spending on markups.

In summary, this Approach 3 outlines an estimation procedure for a model that is strictly more general than that in section 2: it maintains the production function in (2), (3), and (4) and implicitly drops the monopolistic competition assumption by allowing for variable markups. The cost of dropping the monopolistic competition assumption is that additional mean independence restrictions might need to be imposed in order to estimate the parameters determining the elasticity of output with respect to materials, β_m , and with respect to labor, $(\beta_l, \beta_{ll}, \beta_{lk})$.

⁷³If we were to assume that the value added of every affiliate j in every period t is measured without error, then va_{ijt-1} would not be endogenous in equation (A.28). The variable va_{ijt} would still be correlated with u_{ijt} through the period- t firm performance shock η_{ijt} .

A.7 Production Function: Translog in Materials

Assume the demand function is defined by (1) and the production function generalizes that in (2) to be translog in labor, capital, and materials:

$$Q_{ijt} = (H(K_{ijt}, L_{ijt}, M_{ijt}^*; \alpha)) \exp(\omega_{ijt}), \quad (\text{A.31})$$

where

$$H(K_{ijt}, L_{ijt}, M_{ijt}^*; \alpha) \equiv \exp(h(K_{ijt}, L_{ijt}, M_{ijt}^*; \alpha)) \quad (\text{A.32})$$

$$\begin{aligned} h(k_{ijt}, l_{ijt}, m_{ijt}^*; \alpha) &\equiv \alpha_l l_{ijt} + \alpha_k k_{ijt} + \alpha_{ll} l_{ijt}^2 + \alpha_{kk} k_{ijt}^2 + \alpha_{lk} l_{ijt} k_{ijt} \\ &\quad + \beta_m m_{ijt}^* + \beta_{mm} m_{ijt}^{*2} + \beta_{lm} l_{ijt} m_{ijt}^* + \beta_{km} k_{ijt} m_{ijt}^*. \end{aligned} \quad (\text{A.33})$$

Here we use M_{ijt}^* to denote the total expenditure on materials. By definition, $M_{ijt}^* = Y_{ijt}^* - VA_{ijt}^*$, where Y_{ijt}^* and VA_{ijt}^* denote sales revenue and value added, respectively, for affiliate j of firm i in period t . Thus, allowing for measurement error in both sales revenue, $Y_{ijt} = Y_{ijt}^* \exp(\varepsilon_{ijt}^y)$, and value added, $VA_{ijt} = VA_{ijt}^* \exp(\varepsilon_{ijt}^{va})$, we can write the log of measured materials used by affiliate j of firm i in period t as

$$m_{ijt} = y_{ijt} - va_{ijt} = y_{ijt}^* - va_{ijt}^* + \varepsilon_{ijt}^y - \varepsilon_{ijt}^{va} = m_{ijt}^* + \varepsilon_{ijt}^y - \varepsilon_{ijt}^{va}. \quad (\text{A.34})$$

In this case, the revenue function becomes

$$y_{ijt} = \kappa_{n_{ijt}} + h(l_{ijt}, k_{ijt}, m_{ijt}^*; \beta) + \psi_{ijt} + \varepsilon_{ijt}^y, \quad (\text{A.35})$$

where $\kappa_{n_{ijt}} = (1/\sigma)q_{n_{ijt}} + p_{n_{ijt}}$, $\beta = \alpha\iota$, $\beta_m = \alpha_m\iota$, $\psi_{ijt} = \iota(\omega_{ijt} + \xi_{ijt})$ and $\iota = (\sigma - 1)/\sigma$. Introducing our measure of materials in this expression, it becomes

$$\begin{aligned} y_{ijt} &= \kappa_{n_{ijt}} + h(l_{ijt}, k_{ijt}, m_{ijt} - \varepsilon_{ijt}^y + \varepsilon_{ijt}^{va}; \beta) + \psi_{ijt} + \varepsilon_{ijt}^y, \\ &= \kappa_{n_{ijt}} + h(l_{ijt}, k_{ijt}, y_{ijt} - va_{ijt} - \varepsilon_{ijt}^y + \varepsilon_{ijt}^{va}; \beta) + \psi_{ijt} + \varepsilon_{ijt}^y. \end{aligned} \quad (\text{A.36})$$

Combining (8), (9), and (A.36) we obtain

$$\begin{aligned} y_{ijt} &= h(l_{ijt}, k_{ijt}, y_{ijt} - va_{ijt} - \varepsilon_{ijt}^y + \varepsilon_{ijt}^{va}; \beta) \\ &\quad + \rho(y_{ijt-1} - h(l_{ijt-1}, k_{ijt-1}, y_{ijt-1} - va_{ijt-1} - \varepsilon_{ijt-1}^y + \varepsilon_{ijt-1}^{va}; \beta)) \\ &\quad + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \gamma_{n_{ijt}} + u_{ijt}, \end{aligned} \quad (\text{A.37})$$

where the error term is $u_{ijt} = \eta_{ijt} + \varepsilon_{ijt}^y - \rho\varepsilon_{ijt-1}^y$, and $\gamma_{n_{ijt}} = \mu_{n_{ijt}} + \kappa_{n_{ijt}} - \rho\kappa_{n_{ijt-1}}$.

Using (A.37) as estimating equation requires addressing several identification challenges. First, as static (flexible) inputs, labor and materials hired by firm i 's affiliate j during period t are determined after the period- t shock to productivity η_{ijt} is observed by firm i , giving rise to a correlation between both l_{ijt} and $y_{ijt} - va_{ijt}$ and u_{ijt} . Second, measurement error ε_{ijt-1}^y in sales revenue y_{ijt-1} also appears in the error term u_{ijt} in (13) above, giving rise to a correlation between y_{ijt-1} and u_{ijt} . Third, our measure of materials depends on sales revenue from (A.34); therefore, measurement error in sales revenue ε_{ijt}^y will also affect our measure of materials, generating a correlation between u_{ijt} and $y_{ijt} - va_{ijt}$. Fourth, the function $h(\cdot)$ is not linear in materials and, therefore, the measurement error terms in sales ($\varepsilon_{ijt}^y, \varepsilon_{ijt-1}^y$) and value added ($\varepsilon_{ijt}^{va}, \varepsilon_{ijt-1}^{va}$) enters nonlinearly in (A.37). Estimating the parameter vector $(\beta, \rho, \mu_a, \mu_p, \mu_{ap})$ may be attempted following two steps analogous to those described in section 4.1. We describe here several challenges

posed by this approach.

Step 1. Given the production function in (A.31), (A.32), and (A.33), the profit function in (12), and the assumptions that both labor and materials are static inputs, necessary conditions for observed labor and materials to be optimally determined by firm i for its affiliate j are

$$\beta_l + \beta_{ll}2l_{ijt} + \beta_{lk}k_{ijt} + \beta_{lm}(y_{ijt} - m_{ijt} - \varepsilon_{ijt}^y + \varepsilon_{ijt}^{va}) = \frac{W_{ijt}^l}{Y_{ijt}} \exp(\varepsilon_{ijt}^y), \quad (\text{A.38})$$

and

$$\beta_m + \beta_{mm}2(y_{ijt} - m_{ijt} - \varepsilon_{ijt}^y + \varepsilon_{ijt}^{va}) + \beta_{mk}k_{ijt} + \beta_{lm}l_{ijt} = \frac{Y_{ijt} \exp(-\varepsilon_{ijt}^y) - V A_{ijt} \exp(-\varepsilon_{ijt}^{va})}{Y_{ijt} \exp(-\varepsilon_{ijt}^y)}. \quad (\text{A.39})$$

Taking logs on both sides of these two expressions, we obtain

$$\ln(\beta_l + \beta_{ll}2l_{ijt} + \beta_{lk}k_{ijt} + \beta_{lm}(y_{ijt} - m_{ijt} - \varepsilon_{ijt}^y + \varepsilon_{ijt}^{va})) = w_{ijt}^l - y_{ijt} + \varepsilon_{ijt}^y, \quad (\text{A.40})$$

and

$$\begin{aligned} \ln(\beta_m + \beta_{mm}2(y_{ijt} - m_{ijt} - \varepsilon_{ijt}^y + \varepsilon_{ijt}^{va}) + \beta_{mk}k_{ijt} + \beta_{lm}l_{ijt}) \\ = \ln(Y_{ijt} \exp(-\varepsilon_{ijt}^y) - V A_{ijt} \exp(-\varepsilon_{ijt}^{va})) - y_{ijt} + \varepsilon_{ijt}^y. \end{aligned} \quad (\text{A.41})$$

It is not possible to derive from (A.40) and (A.41) moment conditions analogous to those in (15). The reason is that the measurement errors affecting revenue and value added, ε_{ijt}^y and ε_{ijt}^{va} , enter nonlinearly on both the left and right sides of (A.40) and (A.41). Therefore, imposing mean independence restrictions on the distribution of these measurement error terms is not enough for identification. Assuming that both sales revenue and materials are measured without error (i.e. $\varepsilon_{ijt}^y = \varepsilon_{ijt}^{va} = 0$ for all i, j , and t) would not be a solution either: the error terms would also automatically disappear from the right hand side of both (A.40) and (A.41), which would have no unobserved component anymore and, consequently, would be rejected in the data. One solution would be to instead impose parametric assumptions on the distribution of ε_{ijt}^y and ε_{ijt}^{va} , and to accept the consequence that estimates of $(\beta_l, \beta_{ll}, \beta_{lk}, \beta_{lm}, \beta_m, \beta_{mm}, \beta_{mk})$ would be sensitive to such parametric assumptions.

Step 2. Given the conclusion above, we must estimate all parameters of interest using (A.37). The terms ε_{ijt}^y and ε_{ijt}^{va} also enter nonlinearly in the function $h(\cdot)$ and, therefore, in (A.37). Therefore, as discussed in *Step 1* above, a necessary step to use this equation for identification is to assume away measurement error in sales revenue and value added. The resulting expression is:

$$\begin{aligned} y_{ijt} = h(l_{ijt}, k_{ijt}, y_{ijt} - va_{ijt}; \beta) + \rho(y_{ijt-1} - h(l_{ijt-1}, k_{ijt-1}, y_{ijt-1} - va_{ijt-1}; \beta)) \\ + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \gamma_{n_{ijt}} + u_{ijt}, \end{aligned} \quad (\text{A.42})$$

where $u_{ijt} = \eta_{ijt}$, and $\gamma_{n_{ijt}} = \mu_{n_{ijt}} + \kappa_{n_{ijt}} - \rho \kappa_{n_{ijt-1}}$. The challenges that we would need to face to be able to use this equation for estimation is to find instruments for l_{ijt} and $y_{ijt} - va_{ijt}$. Regarding the latter, the ideal instrument would be the price of materials, which we unfortunately do not observe. An alternative would be to use $y_{ijt-2} - va_{ijt-2}$, though this is unlikely to be correlated with $y_{ijt} - va_{ijt}$ after controlling for $y_{ijt-1} - va_{ijt-1}$.

A.8 Nonlinear Evolution of Firm Performance

Assume firms face the demand function defined by (1), the production function defined in (2), and are monopolistically competitive. This yields the revenue equation in (6). Suppose we were to generalize the stochastic process of firm performance so that the expected value of period- t firm performance conditional on the information set of firm i at $t-1$ is allowed to depend on period- $t-1$ performance in a nonlinear way. Specifically, instead of (9), assume that

$$\mathbb{E}_{t-1}[\psi_{ijt}] = \rho_1 \psi_{ijt-1} + \rho_2 \psi_{ijt-1}^2 + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \mu_{n_{ijt}}. \quad (\text{A.43})$$

Combining (6), (8), and (A.43), we obtain

$$\begin{aligned} va_{ijt} = & h(k_{ijt}, l_{ijt}; \beta) + \rho_1 (va_{ijt-1} - h(k_{ijt-1}, l_{ijt-1}; \beta)) + \rho_2 (va_{ijt-1} - h(k_{ijt-1}, l_{ijt-1}; \beta) - \kappa_{n_{ijt-1}})^2 \\ & + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \gamma_{n_{ijt}} + u_{ijt}, \end{aligned} \quad (\text{A.44})$$

where $u_{ijt} = \eta_{ijt} + \varepsilon_{ijt} - \rho \varepsilon_{ijt-1}$, and $\gamma_{n_{ijt}} = \mu_{n_{ijt}} + \kappa_{n_{ijt}} - \rho \kappa_{n_{ijt-1}}$. As it is clear from (A.44), using this equation to estimate the parameter vector of interest, $(\beta, \rho_1, \rho_2, \mu_a, \mu_p, \mu_{ap}, \{\gamma_{n_{ijt}}\}, \{\kappa_{n_{ijt}}\})$, requires estimating also a large set of market-year fixed effects $\{\kappa_{n_{ijt-1}}\}$ that enter nonlinearly in the estimating equation. The estimates of the parameter vector of interest will therefore suffer from asymptotic bias due to an incidental parameters problem.

A.9 NLS and GMM Estimation: Details

For any variable x_{ijt} , use x'_{ijt} to denote the residual from projecting x_{ijt} on a full set of market-year fixed effects, $\{\gamma_{n_{ijt}}\}$. Therefore, from (16), we can write

$$\begin{aligned} \eta'_{ijt} = & \widehat{va}'_{ijt} - \beta_k k'_{ijt} - \beta_{kk} (k_{ijt}^2)' - \rho (\widehat{va}'_{ijt-1} - \beta_k k'_{ijt-1} - \beta_{kk} (k_{ijt-1}^2)') \\ & - \mu_a r'_{ijt-1} - \mu_p r'_{i0t-1} - \mu_{ap} (r_{ijt-1} r_{i0t-1})', \end{aligned} \quad (\text{A.45})$$

where $(k_{ijt}^2)'$, $(k_{ijt-1}^2)'$ and $(r_{ijt-1} r_{i0t-1})'$ denote the residuals from the projection of the variables k_{ijt}^2 , k_{ijt-1}^2 and $r_{ijt-1} r_{i0t-1}$, respectively, on a full set of market-year fixed effects $\{\gamma_{n_{ijt}}\}$.

Using this notation, we can write the NLS estimator for the parameter vector $(\beta_k, \beta_{kk}, \rho, \mu_a, \mu_p, \mu_{ap})$ described in section 4.1 as

$$(\hat{\beta}_k, \hat{\beta}_{kk}, \hat{\rho}, \hat{\mu}_a, \hat{\mu}_p, \hat{\mu}_{ap}) = \min_{(\beta_k, \beta_{kk}, \rho, \mu_a, \mu_p, \mu_{ap})} \sum_{i,j,t} \{ \mathbb{1}\{j \in \mathcal{J}_{it}, j \in \mathcal{J}_{it-1}\} \times \eta'_{ijt} \}, \quad (\text{A.46})$$

where (A.45) implicitly defines η'_{ijt} as a function of the parameters of interest $(\beta_k, \beta_{kk}, \rho, \mu_a, \mu_p, \mu_{ap})$.

Similarly, we can use (A.45) to define the GMM estimator for the parameter vector $(\beta_k, \beta_{kk}, \rho, \mu_a, \mu_p, \mu_{ap})$ described in section 9 and in Appendix A.15. Specifically, we compute the optimal two-step GMM estimator based on the following set of unconditional moments

$$\mathbb{E} \left[\eta'_{ijt} \times \begin{pmatrix} k'_{ijt} \\ (k_{ijt}^2)' \\ k'_{ijt-1} \\ (k_{ijt-1}^2)' \\ \text{UCRD}'_{it-1} \\ (\text{UCRD}_{it-1} \times \text{IPR}_{n_{ijt-1}})' \end{pmatrix} \times \mathbb{1}\{j \in \mathcal{J}_{it}, j \in \mathcal{J}_{it-1}\} \right] = 0,$$

where UCRD'_{it-1} and $(\text{UCRD}_{it-1} \times \text{IPR}_{n_{ijt}-1})'$ denote the residuals from the projection of the variables UCRD_{it-1} and $\text{UCRD}_{it-1} \times \text{IPR}_{n_{ijt}-1}$, respectively, on a full set of market-year fixed effects $\{\gamma_{n_{ijt}}\}$.

A.10 Headquarters Innovation and Affiliate Performance

To evaluate the contribution of firm- i parent innovation to the long-run performance of its affiliate j , we use information on the levels of innovation, r_{i0t} and r_{ijt} , that prevail in the firm during a base period t . Supposing these base-year levels are held constant, the expected long-run performance of j is

$$\begin{aligned}\psi_{ij} &\equiv \mathbb{E}\left[\lim_{s \rightarrow \infty} \psi_{ijs} | r_{i0t}, r_{ijt}\right] \\ &= \sum_{s>t} \rho^{s-t} \mu_{n_{ijs}} + \frac{1}{1-\rho} g(r_{ijt}, r_{i0t}),\end{aligned}\tag{A.47}$$

where r_{i0t} and r_{ijt} are observable, and $g(r_{ijt}, r_{i0t}) \equiv \mathbb{E}_{t-1}[\psi_{ijt}] - \rho\psi_{ijt-1} - \mu_{n_{ijt}}$ with $\mathbb{E}_{t-1}[\psi_{ijt}]$ defined in (9). In order to derive (A.47) we have applied $\mathbb{E}[\eta_{ijs} | r_{ijt}, r_{i0t}] = 0$ for all $s > t$, as implied by (8).

The long-run performance of affiliate j of firm i in the case in which parent R&D is zero and affiliate R&D remains at its period- t level yields

$$\psi_{ij,r_0} = \sum_{s>t} \rho^{s-t} \mu_{n_{ijs}} + \frac{1}{1-\rho} g(r_{ijt}, 0),$$

and we assess the contribution of parent innovation by comparing the distributions of ψ_{ij} and ψ_{ij,r_0} across multinational firm affiliates. Similarly, the long-run performance of affiliate j of firm i in the case in which affiliate R&D is zero and parent R&D remains at its period- t level yields

$$\psi_{ij,r_j} = \sum_{s>t} \rho^{s-t} \mu_{n_{ijs}} + \frac{1}{1-\rho} g(0, r_{i0t}).$$

Note that the difference between any of the terms ψ_{ij} , ψ_{ij,r_j} , and ψ_{ij,r_0} does not depend on the set of fixed effects $\{\gamma_{n_{ijt}}\}$.

A.11 Innovation and the Headquarters Performance Advantage

To compute the long-run performance ψ_{ij} of affiliates j of a multinational i , we use the expression in (A.47) above. From (18), the expected long-run productivity the parent firm of multinational i is

$$\psi_{i0} = \mathbb{E}\left[\lim_{s \rightarrow \infty} \psi_{i0s} | r_{i0t}\right] = \sum_{s>t} \rho^{s-t} \mu_{n_{i0s}} + \frac{1}{1-\rho_0} (\mu_0 r_{i0t} + \mu_{a0} \sum_j r_{ijt-1}).\tag{A.48}$$

Comparing equations (A.47) and (A.48), one can see that the difference in performance between parents and affiliates will depend both on the market-year unobserved exogenous factors that affect the evolution of the performance of the parent, $\mu_{n_{i0t}}$, as well as those unobserved factors that affect the evolution of performance for each affiliate j , $\mu_{n_{ijt}}$. Being able to identify these parameters would require data on the price index, $P_{n_{ijt}}$, the quantity index, $Q_{n_{ijt}}$, and the price of materials, $P_{n_{ijt}}^m$, in every market and year in which either the parent or an affiliate operates. Such data is not available to us; therefore, Figure 2 reports the distribution of the performance of every affiliate relative to its parent that is exclusively due to the distribution of R&D spending within the multinational firm. Specifically, it provides the value of each

percentile of the distribution of

$$\frac{\mu_a r_{ijt} + \mu_p r_{i0t} + \mu_{ap} r_{ijt-1} r_{i0t}}{1 - \rho} - \frac{\mu_0 r_{i0t}}{1 - \rho_0},$$

for $t = 2004$.

A.12 Innovation Policy Effectiveness and the R&D Return

Here we derive the expression for the gross returns to parent R&D investment under the assumption that the number of affiliates and, for each affiliate, their value added, R&D spending and imports from parent remain constant at period- t levels; i.e. equation (20) in the main text.

First, taking into account that the R&D investment performed by the parent at period t only affects the future value added of affiliate j through its impact on period $t + 1$ performance, $\Psi_{ijt+1} = \exp(\psi_{ijt+1})$, we can rewrite the gross return term GR_{i0t} in (19) as

$$\begin{aligned} GR_{i0t} &= \mathbb{E}_t \left[\sum_{s>t} \sum_{j \in \mathcal{J}_{is}} \frac{\partial V A_{ijs}^*}{\partial R_{i0t}} \right] \\ &= \mathbb{E}_t \left[\sum_{s>t} \sum_{j \in \mathcal{J}_{is}} \frac{\partial \Psi_{ijt+1}}{\partial R_{i0t}} \frac{\partial V A_{ijs}^*}{\partial \Psi_{ijt+1}} \right] \\ &= \mathbb{E}_t \left[\sum_{s>t} \sum_{j \in \mathcal{J}_{is}} \frac{\partial \Psi_{ijt+1}}{\partial R_{i0t}} \frac{\partial \Psi_{ijs}}{\partial \Psi_{ijt+1}} \frac{\partial V A_{ijs}^*}{\partial \Psi_{ijs}} \right] \\ &= \mathbb{E}_t \left[\sum_{s>t} \left[\frac{\partial \Psi_{i0t+1}}{\partial R_{i0t}} \frac{\partial \Psi_{i0s}}{\partial \Psi_{i0t+1}} \frac{\partial V A_{i0s}^*}{\partial \Psi_{i0s}} + \sum_{\substack{j \in \mathcal{J}_{is} \\ j \neq 0}} \frac{\partial \Psi_{ijt+1}}{\partial R_{i0t}} \frac{\partial \Psi_{ijs}}{\partial \Psi_{ijt+1}} \frac{\partial V A_{ijs}^*}{\partial \Psi_{ijs}} \right] \right] \\ &= \mathbb{E}_t \left[\sum_{s>t} \left[\frac{\partial \psi_{i0t+1}}{\partial r_{i0t}} \rho_0^{s-t-1} \frac{V A_{i0s}^*}{R_{i0t}} + \sum_{\substack{j \in \mathcal{J}_{is} \\ j \neq 0}} \frac{\partial \psi_{ijt+1}}{\partial r_{i0t}} \rho^{s-t-1} \frac{V A_{ijs}^*}{R_{i0t}} \right] \right], \end{aligned} \quad (\text{A.49})$$

where the second equality applies the chain rule, the third equality differentiates between the impact of parent R&D on the parent itself and all its affiliates, and the fourth equality uses the fact that

$$\begin{aligned} \frac{\partial \Psi_{ijt+1}}{\partial R_{i0t}} &= \frac{\partial \psi_{ijt+1}}{\partial r_{i0t}} \frac{\Psi_{ijt+1}}{R_{i0t}}, & j = 0, \dots, J_{is}, \\ \frac{\partial \Psi_{i0s}}{\partial \Psi_{i0t+1}} &= \rho_0^{s-t-1} \frac{\Psi_{i0s}}{\Psi_{i0t+1}}, \\ \frac{\partial \Psi_{ijs}}{\partial \Psi_{ijt+1}} &= \rho^{s-t-1} \frac{\Psi_{ijs}}{\Psi_{ijt+1}}, & j = 1, \dots, J_{is}, \\ \frac{\partial V A_{ijs}^*}{\partial \Psi_{ijs}} &= \frac{V A_{ijs}^*}{\Psi_{ijs}}, & j = 0, \dots, J_{is}. \end{aligned}$$

Furthermore, assuming the specifications of the stochastic process of productivity given by (9) and (18), it will be true that

$$\frac{\partial \psi_{i0t+1}}{\partial r_{i0t}} = \mu_0, \quad (\text{A.50a})$$

$$\frac{\partial \psi_{ijt+1}}{\partial r_{i0t}} = \mu_p + \mu_{ap} r_{ijt}. \quad (\text{A.50b})$$

Plugging these expressions into (A.49), we obtain

$$GR_{i0t} = \mathbb{E}_t \left[\sum_{s>t} \left[\mu_0 \rho_0^{s-t-1} \frac{VA_{i0s}^*}{R_{i0t}} + \sum_{j=1}^{J_{is}} (\mu_p + \mu_{ap} r_{ijt}) \rho^{s-t-1} \frac{VA_{ijs}^*}{R_{i0t}} \right] \right]. \quad (\text{A.51})$$

Finally, assuming that J_{is} , VA_{i0s}^* and VA_{ijs}^* remain constant at their period t values for every year $s > t$

$$GR_{i0t} = \frac{\mu_0}{1 - \rho_0} \frac{VA_{i0t}^*}{R_{i0t}} + \sum_{\substack{j \in J_{it} \\ j \neq 0}} \frac{\mu_p + \mu_{ap} r_{ijt}}{1 - \rho} \frac{VA_{ijt}^*}{R_{i0t}}, \quad (\text{A.52})$$

with corresponds to (20) in the main text.

From (A.51), we can also compute the derivative of GR_{i0t} with respect to R_{i0t} ,

$$\begin{aligned} \frac{\partial GR_{i0t}}{\partial R_{i0t}} = \mathbb{E}_t \left[\sum_{s>t} \left[\frac{\partial \psi_{i0t+1}}{\partial r_{i0t}} \rho_0^{s-t-1} \frac{1}{(R_{i0t})^2} \left(\frac{\partial VA_{i0s}^*}{\partial R_{i0t}} R_{i0t} - VA_{i0s}^* \right) \right. \right. \\ \left. \left. + \sum_{\substack{j \in J_{is} \\ j \neq 0}} \frac{\partial \psi_{ijt+1}}{\partial r_{i0t}} \rho^{s-t-1} \frac{1}{(R_{i0t})^2} \left(\frac{\partial VA_{ijs}^*}{\partial R_{i0t}} R_{i0t} - VA_{ijs}^* \right) \right] \right], \end{aligned}$$

and, therefore,

$$\begin{aligned} \frac{\partial GR_{i0t}}{\partial R_{i0t}} = \mathbb{E}_t \left[\sum_{s>t} \left[\frac{\partial \psi_{i0t+1}}{\partial r_{i0t}} \rho_0^{s-t-1} \frac{VA_{i0s}^*}{(R_{i0t})^2} \left(\frac{\partial VA_{i0s}^*}{\partial R_{i0t}} \frac{R_{i0t}}{VA_{i0s}^*} - 1 \right) \right. \right. \\ \left. \left. + \sum_{\substack{j \in J_{is} \\ j \neq 0}} \frac{\partial \psi_{ijt+1}}{\partial r_{i0t}} \rho^{s-t-1} \frac{VA_{ijs}^*}{(R_{i0t})^2} \left(\frac{\partial VA_{ijs}^*}{\partial R_{i0t}} \frac{R_{i0t}}{VA_{ijs}^*} - 1 \right) \right] \right], \end{aligned}$$

or, equivalently,

$$\frac{\partial GR_{i0t}}{\partial R_{i0t}} = \mathbb{E}_t \left[\sum_{s>t} \left[\frac{\partial \psi_{i0t+1}}{\partial r_{i0t}} \rho_0^{s-t-1} \frac{VA_{i0s}^*}{(R_{i0t})^2} \left(\frac{\partial va_{i0s}^*}{\partial r_{i0t}} - 1 \right) + \sum_{\substack{j \in J_{is} \\ j \neq 0}} \frac{\partial \psi_{ijt+1}}{\partial r_{i0t}} \rho^{s-t-1} \frac{VA_{ijs}^*}{(R_{i0t})^2} \left(\frac{\partial va_{ijs}^*}{\partial r_{i0t}} - 1 \right) \right] \right].$$

Noticing that

$$\begin{aligned} \frac{\partial va_{i0s}^*}{\partial r_{i0t}} &= \frac{\partial \psi_{i0t+1}}{\partial r_{i0t}} \rho_0^{s-t-1}, \\ \frac{\partial va_{ijs}^*}{\partial r_{i0t}} &= \frac{\partial \psi_{ijt+1}}{\partial r_{i0t}} \rho^{s-t-1}, \end{aligned}$$

and using (A.50a) and (A.50b), the expression simplifies to

$$\begin{aligned} \frac{\partial GR_{i0t}}{\partial R_{i0t}} = \mathbb{E}_t \left[\sum_{s>t} \left[\mu_0 \rho_0^{s-t-1} \frac{VA_{i0s}^*}{(R_{i0t})^2} \left(\mu_0 \rho_0^{s-t-1} - 1 \right) + \sum_{\substack{j \in J_{is} \\ j \neq 0}} (\mu_p + \mu_{ap} r_{ijt}) \rho^{s-t-1} \frac{VA_{ijs}^*}{(R_{i0t})^2} \right. \right. \\ \left. \left. \times \left((\mu_p + \mu_{ap} r_{ijt}) \rho^{s-t-1} - 1 \right) \right] \right]. \end{aligned}$$

Notice that the derivative above depends on the number and size of affiliates, and their involvement in R&D performance.

A.13 U.S. Parent Innovation and GDP Growth Abroad

Denote the aggregate value added of all firms operating in market n at period t as VA_{nt}^* , and the aggregate value added of all affiliates of U.S.-based multinationals operating in market n at period t as $VA_{US,nt}^*$. Then, $VA_{US,nt}^* = \sum_{j \in \mathcal{J}_{nt}} VA_{ijt}^*$, where \mathcal{J}_{nt} denotes the set of affiliates of U.S. multinationals operating in market n in period t . Similarly, denote as VA'_{ijt} the counterfactual value added of affiliate j of multinational firm i in period t if the only change in the environment is that it cannot benefit from the R&D investment performed by its parent at $t-1$ (i.e. every other determinant of value added is kept at its observed period- t value). Similarly, let's define $VA'_{US,nt} = \sum_{j \in \mathcal{J}_{nt}} VA'_{ijt}$ and $\Delta_{US,nt} = VA'_{US,nt}/VA_{US,nt}^*$. Then, using the expression for the evolution of affiliate productivity in (9),

$$\Delta_{US,nt} = \frac{\sum_{j \in \mathcal{J}_{nt}} \exp(va_{ijt} - \varepsilon_{ijt} - \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1})}{\sum_{j \in \mathcal{J}_{nt}} \exp(va_{ijt} - \varepsilon_{ijt})}.$$

Notice that the estimation procedure in section 4.1 provides estimates for all the parameters entering this expression, including the measurement error term ε_{ijt} . Once we have computed the value of $\Delta_{US,nt}$, we can measure the relative change in VA_{nt}^* due to affiliates of U.S. parents not being able to benefit from their R&D investment. Specifically, denoting VA'_{nt} as the total value added generated by all firms in market n at period t in the counterfactual scenario of interest, and $VA_{H,nt}^*$ as the total value added generated by non-U.S. affiliates located in market n at period t , we may write

$$\Delta_{nt} = \frac{VA'_{nt}}{VA_{nt}^*} = \frac{VA'_{US,nt} + VA_{H,nt}^*}{VA_{nt}^*} = \frac{\Delta_{US,nt} \sum_{j \in \mathcal{J}_{nt}} \exp(va_{ijt} - \varepsilon_{ijt}) + (VA_{nt}^* - \sum_{j \in \mathcal{J}_{nt}} \exp(va_{ijt} - \varepsilon_{ijt}))}{VA_{nt}^*}.$$

To measure VA_{nt}^* as the GDP of market n at period t , as reported in the database INDSTAT4; i.e. UNIDO Industrial Statistics Database at the 4-digit ISIC level.

A.14 Intrafirm Trade: Model and Empirical Strategy

We extend here the model presented in section 2 to allow affiliates to source inputs from their parents. Specifically, we generalize the affiliates' production function in equation (2) and assume that, to produce output Q_{ijt} , affiliate j combines capital, labor, materials sourced from third-party supplies and materials sourced from their parent using the following production technology

$$Q_{ijt} = (H(K_{ijt}, L_{ijt}; \alpha))^{1-\alpha_m-\alpha_{im}} M_{ijt}^{\alpha_m} IM_{ijt}^{\alpha_{im}} \exp(\omega_{ijt}) \quad (\text{A.53})$$

where the function $H(\cdot)$ is defined as in equations (3) and (4), K_{ijt} is effective units of capital, L_{ijt} is the number of production workers, M_{ijt} is an unobserved quantity index of materials use sourced from third parties, IM_{ijt} is effective units of materials sourced by affiliate j from its parent, and ω_{ijt} denotes the Hicks-neutral physical productivity at t . The optimal amount of materials that affiliate j buys from either third party suppliers or from its parent are

$$\begin{aligned} \max_{M_{ijt}, IM_{ijt}} \{Y_{ijt} - P_{nijt}^m M_{ijt} - P_{nijt}^{im} IM_{ijt}\} = \\ \max_{M_{ijt}, IM_{ijt}} \{P_{ijt} Q_{ijt} - P_{nijt}^m M_{ijt} - P_{nijt}^{im} IM_{ijt}\} = \\ \max_{M_{ijt}, IM_{ijt}} \{Q_{nijt}^{\frac{1}{\sigma}} P_{nijt} \exp\left(\xi_{ijt} \frac{\sigma-1}{\sigma}\right) Q_{ijt}^{\frac{\sigma-1}{\sigma}} - P_{nijt}^m M_{ijt} - P_{nijt}^{im} IM_{ijt}\}, \end{aligned}$$

which is identical to the maximization problem described in Appendix (A.1) except for the fact that now Q_{ijt} stands for the production function in equation (A.53) instead of that in equation (2) and that now we are optimizing over both materials coming from outside the firm M_{ijt} and materials coming from the parent firm IM_{ijt} . Given the production function in (A.53), the optimal level of material inputs M_{ijt}^* and IM_{ijt}^* satisfy the following two conditions

$$P_{n_{ijt}}^m = Q_{n_{ijt}}^{\frac{1}{\sigma}} P_{n_{ijt}} \exp \left[\frac{(\xi_{ijt} + \omega_{ijt})(\sigma - 1)}{\sigma} \right] [H(K_{ijt}, L_{ijt}; \alpha)]^{\frac{(1 - \alpha_m - \alpha_{im})(\sigma - 1)}{\sigma}} \times \frac{\alpha_m(\sigma - 1)}{\sigma} (M_{ijt}^*)^{\frac{\alpha_m(\sigma - 1)}{\sigma} - 1} (IM_{ijt}^*)^{\frac{\alpha_{im}(\sigma - 1)}{\sigma}},$$

and

$$P_{n_{ijt}}^{im} = Q_{n_{ijt}}^{\frac{1}{\sigma}} P_{n_{ijt}} \exp \left[\frac{(\xi_{ijt} + \omega_{ijt})(\sigma - 1)}{\sigma} \right] [H(K_{ijt}, L_{ijt}; \alpha)]^{\frac{(1 - \alpha_m - \alpha_{im})(\sigma - 1)}{\sigma}} \times \frac{\alpha_{im}(\sigma - 1)}{\sigma} (M_{ijt}^*)^{\frac{\alpha_m(\sigma - 1)}{\sigma}} (IM_{ijt}^*)^{\frac{\alpha_{im}(\sigma - 1)}{\sigma} - 1}.$$

Solving for both M_{ijt}^* and IM_{ijt}^* , we obtain

$$\frac{M_{ijt}^*}{IM_{ijt}^*} = \frac{\alpha_m}{\alpha_{im}} \frac{P_{n_{ijt}}^{im}}{P_{n_{ijt}}^m}$$

and, following the same steps as in Appendix A.1, we can rewrite the revenue function as

$$y_{ijt} = \tilde{\kappa}'_{n_{ijt}} + h(k_{ijt}, l_{ijt}; \beta') + \psi_{ijt},$$

where $\tilde{\kappa}'_{n_{ijt}}$ is a function of the vector $(\alpha, \alpha_m, \alpha_{im}, \sigma, p_{n_{ijt}t}, q_{n_{ijt}t}, p_{n_{ijt}t}^m, p_{n_{ijt}t}^{im})$ and $\beta' \equiv \iota' \alpha$ where ι is a function of α_m, α_{im} and σ . Obtaining a similar expression for the value added function is also straightforward. The first order condition for both materials inputs is

$$P_{n_{ijt}}^m M_{ijt}^* = \left(\frac{\alpha_m(\sigma - 1)}{\sigma} \right) Y_{ijt}, \quad (\text{A.54})$$

$$P_{n_{ijt}}^{im} IM_{ijt}^* = \left(\frac{\alpha_{im}(\sigma - 1)}{\sigma} \right) Y_{ijt}, \quad (\text{A.55})$$

which implies that value added (in logs) may thus be concisely represented as,

$$va_{ijt}^* = \kappa'_{n_{ijt}} + h(k_{ijt}, l_{ijt}; \beta') + \psi_{ijt}, \quad (\text{A.56})$$

where

$$\kappa'_{n_{ijt}} = \ln \left(1 - \frac{\alpha_m(\sigma - 1)}{\sigma} - \frac{\alpha_{im}(\sigma - 1)}{\sigma} \right) + \tilde{\kappa}'_{n_{ijt}}.$$

Equation (A.56) is analogous to equation (5). The only differences between these two equations are the exact functions of the structural parameters that the market-year unobserved effects, $\kappa'_{n_{ijt}}$, and the value-added elasticities, β' , represent.

This implies that assuming that the production of each affiliate j at t is as represented in equation (A.53) instead of as represented in equation (2) would not have any impact on the estimating equation in equation (13) nor on the estimation procedure described in Section 4. The results presented in Section 5 are therefore consistent with a model in which each affiliate's demand is defined by equation (1), each affiliates production is defined by equation (A.53) and the evolution of productivity is defined as in equations (8) and

(9). The key condition so that the results presented in Section 5 generalize to a model in which affiliates' production function is as represented in equation (A.53) is that affiliates decide how much material inputs to source from their parent taking the price set by this parent as given and that this price is constant for all affiliates located in the same market-year.

The key difference between the model estimated in Section 8 and the horizontal model discussed in the previous sections is therefore the replacement of equation (9) by (22). This implies that the estimating equation for the vertical model in Section 8 becomes a trivial extension of that in equation (13). Specifically, the new estimating equation becomes:

$$\begin{aligned} va_{ijt} = & h(k_{ijt}, l_{ijt}; \beta) + \rho(va_{ijt-1} - h(k_{ijt-1}, l_{ijt-1}; \beta)) + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} \\ & + \mu_{ap} r_{ijt-1} r_{i0t-1} + \mu_{pm} r_{i0t-1} im_{ijt-1} + \mu_{apm} r_{ijt-1} r_{i0t-1} im_{ijt-1} + \mu_m im_{ijt-1} + \gamma'_{n_{ijt}} + u_{ijt}, \end{aligned} \quad (\text{A.57})$$

where $u_{ijt} \equiv \eta_{ijt} + \varepsilon_{ijt} - \rho \varepsilon_{ijt-1}$ is a function of the performance shock and measurement error in value added, and where $\gamma'_{n_{ijt}} \equiv \mu_{n_{ijt}} + \kappa'_{n_{ijt}} - \rho \kappa'_{n_{ijt-1}}$ is a market-year effect that accounts for both the unobserved quantity and prices embedded in $\kappa'_{n_{ijt}}$ and the market-year specific component of firm performance, $\mu_{n_{ijt}}$. Furthermore, as $\mathbb{E}[u_{ijt}|im_{ijt-1}] = 0$, the discussion of the estimation procedure and identification concerns in Section 4 also applies to the model with intrafirm imports discussed in Section 8.

A.15 Misreporting of R&D Expenditures

Measurement error in R&D expenditures may affect parameters estimated in Step 2 of the procedure in section 4.1. Specifically, using the notation introduced in Section 9, we define

$$r_{ijt-1} = r_{ijt-1}^* + x_{ijt-1}, \quad \forall j = 0, \dots, \mathcal{J}_{it},$$

where r_{ijt-1}^* denotes the true investment in R&D of site j of firm i at $t-1$, r_{ijt-1} denotes the reported investment and x_{ijt-1} denotes the degree of over-reporting in R&D spending. Using this notation, we can rewrite the estimating equation in (16) as

$$\begin{aligned} \widehat{va}_{ijt} = & \beta_k k_{ijt} + \beta_{kk} k_{ijt}^2 + \rho(\widehat{va}_{ijt-1} - \beta_k k_{ijt-1} - \beta_{kk} k_{ijt-1}^2) + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} \\ & + \mu_{ap} r_{ijt-1} r_{i0t-1} + \gamma_{n_{ijt}} + \nu_{ijt}, \end{aligned} \quad (\text{A.58})$$

where the error term is

$$\nu_{ijt} = \eta_{ijt} - \mu_a x_{ijt-1} - \mu_p x_{i0t-1} - \mu_{ap} x_{ijt-1} x_{i0t-1}. \quad (\text{A.59})$$

For clarity in the exposition, assume $\mu_p = \mu_{ap} = 0$. In this case, we can rewrite the NLS estimate of μ_a as

$$\frac{cov_c(\widehat{va}_{ijt}, r_{ijt-1})}{var_c(r_{ijt-1})} = \frac{cov_c(\mu_a r_{ijt-1} + \nu_{ijt}, r_{ijt-1})}{var_c(r_{ijt-1})} = \frac{cov_c(\mu_a r_{ijt-1} + \eta_{ijt} - \mu_a x_{ijt-1}, r_{ijt-1})}{var_c(r_{ijt-1})},$$

where $cov_c(\cdot)$ and $var_c(\cdot)$ denote the covariance and variance after conditioning on the vector $(k_{ijt}, \widehat{va}_{ijt-1}, k_{ijt-1})$ and a full set of country-sector-year fixed effects. Through simple algebra and imposing both the mean independence condition in (8) and assuming that the misreporting error x_{ijt-1} is independent of the performance shock η_{ijt} , we can rewrite this expression as

$$\mu_a + \frac{cov_c(-\mu_a x_{ijt-1}, r_{ijt-1}^* + x_{ijt-1})}{var_c(r_{ijt-1}^* + x_{ijt-1})} = \mu_a - \mu_a \frac{cov_c(x_{ijt-1}, r_{ijt-1}^*) + var_c(x_{ijt-1})}{var_c(r_{ijt-1}^* + x_{ijt-1})}.$$

The asymptotic bias of this estimator would therefore be

$$-\mu_a \frac{\text{cov}_c(x_{ijt-1}, r_{ijt-1}^*) + \text{var}_c(x_{ijt-1})}{\text{var}_c(r_{ijt-1}^* + x_{ijt-1})}.$$

As, by definition, $\text{var}_c(x_{ijt-1}) > 0$ and $\text{var}_c(r_{ijt-1}^* + x_{ijt-1}) > 0$, it is clear that $\text{cov}_c(x_{ijt-1}, r_{ijt-1}^*) \geq 0$ implies a downward bias in our benchmark estimates of the elasticity of period- t affiliate performance with respect to period- $t-1$ affiliate R&D expenditure; i.e. downward bias in the estimate of μ_a computed following the procedure in 4.1. Analogously, we can show that, if the correlation between x_{i0t-1} and r_{i0t-1}^* is weakly positive, the estimation procedure described in Section 4.1 will yield estimates of the elasticity of period- t affiliate performance with respect to period- $t-1$ parent R&D expenditure that will also be asymptotically downward biased.

Why shall we expect the covariance between the over-reporting in R&D expenditures, x_{ijt-1} , and the actual level of R&D expenditures, r_{ijt-1}^* , to be weakly positive? Suppose that the main factor affecting x_{ijt-1} is R&D subsidies granted by the host country of affiliate j of firm i in year $t-1$. Ceteris paribus, the larger these subsidies are, the larger the expected over-reporting of affiliate R&D; i.e. $\text{cov}_c(x_{ijt-1}, \text{subsidies}_{ijt-1}) \geq 0$. Similarly, ceteris paribus, the R&D subsidies granted by the host country of affiliate j of firm i in year $t-1$ will also have a positive impact on the actual amount of R&D investment performed by affiliate j of firm i at t ; i.e. $\text{cov}_c(r_{ijt-1}^*, \text{subsidies}_{ijt-1}) \geq 0$. Therefore, one can expect the covariance between x_{ijt-1} and r_{ijt-1}^* to be positive and our benchmark estimates of μ_a to be asymptotically downward biased.

In order to address the downward bias in our benchmark estimates of μ_a , μ_p and μ_{ap} that will likely arise from misreporting of R&D expenditures, we use $\text{UCRD}_{it} \times \text{IPR}_{n_{ijt}}$ and UCRD_{it} as instruments to form moment conditions that allow us to compute Generalized Method of Moments (GMM) estimates of the parameter vector $(\rho, \beta_k, \beta_{kk}, \mu_a, \mu_p, \mu_{ap})$.

Specifically, maintaining again the assumption that $\mu_p = \mu_{ap} = 0$ for clarity in the exposition, we can rewrite the GMM estimator of μ_a as

$$\frac{\text{cov}_c(\widehat{v}a_{ijt}, \text{UCRD}_{it-1} \times \text{IPR}_{n_{ijt-1}})}{\text{cov}_c(r_{ijt-1}, \text{UCRD}_{it-1} \times \text{IPR}_{n_{ijt-1}})},$$

or, equivalently,

$$\frac{\text{cov}_c(\mu_a r_{ijt-1} + \eta_{ijt} - \mu_a x_{ijt-1}, \text{UCRD}_{it-1} \times \text{IPR}_{n_{ijt-1}})}{\text{cov}_c(r_{ijt-1}, \text{UCRD}_{it-1} \times \text{IPR}_{n_{ijt-1}})}$$

where, as above, $\text{cov}_c(\cdot)$ denotes the covariance after conditioning on the vector $(k_{ijt}, \widehat{v}a_{ijt-1}, k_{ijt-1})$ and a full set of country-sector-year fixed effects. By simple algebra, one can show that, if the covariance between our instrument and both the productivity shock η_{ijt} and the error in the reported R&D expenditure is zero conditional on lagged value added, capital, lagged capital, and a full set of market-year fixed effects, then our GMM estimate of μ_a will be consistent. We argue in section 9 why we think these assumptions are reasonable.

Similarly, assuming that $\mu_a = \mu_{ap} = 0$ for simplicity, we can rewrite our GMM estimation of μ_p as

$$\frac{\text{cov}_c(\widehat{v}a_{ijt}, \text{UCRD}_{it-1})}{\text{cov}_c(r_{i0t-1}, \text{UCRD}_{it-1})} = \frac{\text{cov}_c(\mu_p r_{0jt-1} + \nu_{ijt}, \text{UCRD}_{it-1})}{\text{cov}_c(r_{i0t-1}, \text{UCRD}_{it-1})} = \frac{\text{cov}_c(\mu_p r_{0jt-1} + \eta_{ijt} - \mu_p x_{i0t-1}, \text{UCRD}_{it-1})}{\text{cov}_c(r_{i0t-1}, \text{UCRD}_{it-1})}.$$

If we assume that the user cost of R&D in the U.S. state of the parent site at some period $t-1$ is uncorrelated with the period t productivity shock of one of its affiliates, then we can rewrite the asymptotic bias of our

GMM estimate of μ_p as

$$-\mu_p \frac{\text{cov}_c(x_{i0t-1}, \text{UCRD}_{it-1})}{\text{cov}_c(r_{i0t-1}, \text{UCRD}_{it-1})}.$$

As long as the sign of $\text{cov}_c(x_{i0t-1}, \text{UCRD}_{it-1})$ is the same as that of $\text{cov}_c(r_{i0t-1}, \text{UCRD}_{it-1})$, we expect our GMM estimate of μ_p to be asymptotically downward biased. Specifically, as discussed in Section 9, higher R&D subsidies in the state of location of the parent of firm i will imply: (a) lower user cost of R&D; i.e. lower UCRD_{it-1} ; (b) higher over-reporting of R&D expenditures by the parent; i.e. higher x_{i0t-1} ; (c) higher actual investment in R&D by the parent; i.e. higher r_{i0t-1}^* . We thus expect $\text{cov}_c(x_{i0t-1}, \text{UCRD}_{it-1}) < 0$ and $\text{cov}_c(r_{i0t-1}, \text{UCRD}_{it-1}) < 0$. Under these conditions, our GMM estimate of μ_p is expected to be asymptotically downward biased.

Summing up, in the presence of R&D misreporting by affiliates and parents of multinational firms, we expect our NLS estimates of the elasticity of affiliates' performance with respect to parent and affiliate R&D expenditures to be downward biased. Similarly, we expect the GMM estimator that uses both UCRD_{it-1} and $\text{UCRD}_{it-1} \times \text{IPR}_{n_{ijt-1}}$ as instruments for r_{ijt-1} and r_{i0t-1} to yield asymptotically unbiased estimates of the elasticity of affiliate performance with respect to its own R&D investment and asymptotically downward biased estimates of the elasticity of affiliate performance with respect to its parent R&D investment.

Columns 1 through 6 in Table A.2 show that UCRD_{it} and $\text{UCRD}_{it} \times \text{IPR}_{n_{ijt}}$ are systematically related to parent and affiliate R&D spending. Parent firms located in a U.S. state with a relatively high user cost of R&D UCRD_{it} have, in turn, relatively low levels of R&D spending (columns 3–6). When interacted with the level of intellectual property protection in affiliate host countries, a higher user cost of R&D in the U.S.-parent state is also associated with increased foreign-affiliate R&D spending within the same firm, particularly among affiliates in locations with strong intellectual property protection (columns 1–3). The data are thus consistent with the idea that an increase in the user cost of R&D in the United States leads to a reallocation of R&D investment within multinational firms away from the U.S. parent site and toward firm affiliates in countries with strong intellectual property rights. It is worth noting that the correlation between affiliate R&D investment and $\text{UCRD}_{it} \times \text{IPR}_{n_{ijt}}$ is robust to the inclusion of market-year fixed effects in column 2, as is the correlation between parent R&D and UCRD_{it} in column 6. Nevertheless, the regressions in Table A.2 should not be interpreted as a formal first stage of a Two-Stage Least Squares estimation, and are computed only as a representation of the correlation present in our data between r_{ijt} and r_{i0t} , on one side, and $\text{UCRD}_{it} \times \text{IPR}_{n_{ijt}}$ and UCRD_{it} , on the other side. The model estimated is over-identified and imposes nonlinear restrictions on the coefficients; a simple regression of each endogenous covariate on all exogenous covariates is therefore not a well-defined first stage.

Table A.2: Determinants of Parent and Affiliate R&D

Dependent variable:	Correlations					
	r_{ijt}	r_{ijt}	r_{ijt}	r_{i0t}	r_{i0t}	r_{i0t}
	(1)	(2)	(3)	(4)	(5)	(6)
UCRD $_{it} \times \text{IPR}_{n_{ijt}}$	0.730 ^a (0.056)	4.847 ^b (2.083)	3.889 ^b (1.917)	\cdot \cdot	\cdot \cdot	\cdot \cdot
UCRD $_{it}$	1.014 (1.243)	-9.558 (8.326)	9.371 (8.803)	-4.185 ^a (1.602)	-9.478 ^a (2.264)	-7.750 (6.673)
Observations	4,194	4,194	4,194	536	536	536
R^2	0.128	0.254	0.505	0.437	0.458	0.727
F test statistic	171.1	5.43	3.69	6.81	17.6	1.35

Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. Correlations are OLS estimates. Standard errors are in parentheses and F statistics correspond to the instrument. Included controls are $\ln L_{\text{aff}}$ in columns 1–3 and $\ln L_{\text{par}}$ in columns 4–6. Fixed effects included are market-year (columns 2, 3, 5, and 6) and firm (columns 3 and 6).

Table A.3: Results for Baseline Specification: Parent

	NLS			GMM			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Persistence	0.9180 ^a (0.0166)	0.9195 ^a (0.0170)	0.9147 ^a (0.0171)	0.8891 ^a (0.0293)	0.8948 ^a (0.0290)	0.8465 ^a (0.0341)	0.8109 ^a (0.0527)
Parent R&D	0.0106 ^a (0.0033)	0.0109 ^a (0.0033)	0.0125 ^a (0.0034)	0.0216 ^a (0.0066)	0.0249 ^a (0.0095)	0.0297 ^a (0.0001)	0.0695 ^a (0.0326)
Sum Affiliates R&D		-0.0007 (0.0019)			-0.0033 (0.0057)		-0.0247 (0.0194)
Sum Affiliates R&D -same sector-			-0.0031 (0.0014)				
Sum Affiliates R&D -other sector-			0.0035 (0.0026)				
Labor Elasticity	0.8057 (0.0106)	0.8057 (0.0106)	0.8057 (0.0106)	0.8057 (0.0106)	0.8057 (0.0106)	0.8057 (0.0323)	0.8057 (0.0106)
Capital Elasticity	0.0323 (0.0122)	0.0340 (0.0123)	0.0515 (0.0257)	0.0323 (0.0122)	0.0332 (0.0123)	0.2094 (0.0122)	0.0357 (0.0205)
Observations	536	536	536	536	536	536	536

Notes: *a* denotes 1% significance, *b* denotes 5% significance, *c* denotes 10% significance. Columns (1) to (3) report Nonlinear Least Squares estimates; columns (4) to (7) report optimal two-step Generalized Method of Moments estimators of the same parameters. All columns control for year fixed effects; columns (6) and (7) also control for fixed effects for U.S. state of the firm's U.S. headquarters. Standard errors are reported in parenthesis. Persistence corresponds to estimates of ρ . Labor Elasticity is the average value of $\beta_l + \beta_{ll}2l_{ijt} + \beta_{lk}k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk}2k_{ijt} + \beta_{lk}l_{ijt}$. The standard deviation for each of these input elasticities appears in parentheses below its mean. All other estimates capture the elasticity of period *t* performance with respect to the *t* - 1 value of the corresponding covariate. The instrument for parent R&D is the user cost of R&D, which varies by U.S.-state and year and is available from Wilson (2009); the instrument for the sum of affiliates R&D is the analogous sum of the interaction between (a) the user cost of R&D prevailing in the U.S. state of the affiliate's corresponding parent, and (b) the strength of intellectual property rights in the affiliate country from Ginarte and Park (1997) and Park (2008); this instrument varies across U.S. state-country-year triplets. Measures of labor, capital, value added, and R&D expenditure are from the Bureau of Economic Analysis Survey of U.S. Direct Investment Abroad.