The Impact of Innovation in the Multinational Firm*

L. Kamran Bilir
University of Wisconsin-Madison

Eduardo Morales
Princeton University and NBER

April 2016

Abstract

When firms operate production plants in multiple countries, technological improvements developed in one country may be shared with firm sites abroad for efficiency gain. We develop a dynamic model that allows for such intrafirm transfer, and apply it to measure the impact of innovation on performance for a panel of U.S. multinationals. Our estimates indicate U.S. parent R&D raises performance significantly at firm locations abroad, and also complements R&D by affiliates. Parent R&D is a substantially more important determinant of firm performance than affiliate R&D. We identify these R&D effects using variation in location-specific innovation policies.

*We thank Pol Antràs, Costas Arkolakis, Andy Bernard, Jan De Loecker, Amit Gandhi, Elhanan Helpman, Sam Kortum, Nina Pavcnik, Veronica Rappaport, Joel Rodrigue, Andrés Rodríguez-Clare, Felix Tintelnot, Jim Tybout, David Weinstein, Daniel Xu, Steve Yeaple, Bill Zeile, and seminar participants at AEA, Columbia, Dartmouth, Duke, George Washington University, HBS, LSE, Michigan, Minnesota, NBER, Kellogg, Penn State, Princeton, SED, UCLA, 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. 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. E-mail: bilir@wisc.edu, ecmorale@princeton.edu.

1 Introduction

Multinational corporations account for nearly all private innovation investment in the United States.[1] These firms maintain production affiliates in multiple countries, but concentrate investments in research and development (R&D) at headquarters sites, raising the question of how resulting proprietary technologies—and the productivity gains associated with their use—are distributed across affiliates. Despite the potentially substantial impact of such technologies on countries hosting affiliates, the actual extent to which proprietary technology developed by a multinational firm in one location impacts performance at its affiliate sites abroad is not known.

Consider firms in the computer hardware industry. The multinational firm Western Digital locates its headquarters and primary research labs in the United States, but operates major production affiliates in Asia.[2] These affiliates are potentially influenced by the innovation of their U.S. parent, which determines new product and process designs as well as the quality of intermediate inputs provided to affiliates for further processing.[3] Yet, the extent to which observed U.S. innovation investment by a firm like Western Digital measurably improves its foreign affiliate performance is an open empirical question.

This paper evaluates how affiliate-level performance within the multinational firm responds separately to headquarters and affiliate innovation. Guided by a dynamic model of multinational innovation and production, we estimate the impact of site-specific R&D investment on affiliate performance for a panel of U.S.-based multinationals operating in 60 countries. Our results indicate that headquarters innovation systematically increases affiliate performance within the same firm. Affiliates do not all benefit equally: the elasticity of affiliate performance with respect to parent innovation increases in affiliates' own innovation, and in the volume of imports from the U.S. parent. We find that headquarters innovation is the primary determinant of affiliates' long-run performance. By contrast, R&D by an affiliate affects only its own performance.

In the model, affiliates produce using capital, labor, and material inputs, selling output in monopolistically competitive markets similar to De Loecker (2011). Each affiliate is associated with a distinct performance level, determined by productivity and demand shifters, that evolves according to a Markov process. We allow this process to respond differentially to R&D investments of the affiliate, its U.S. parent, or other affiliates within the same firm. These potentially distinct impacts of parent and affiliate innovation are estimated building on Aw, Roberts, and Xu (2011), Doraszelski and Jaumandreu (2013), and Gandhi, Navarro, and Rivers (2013).

To estimate the unknown parameters of the model, we use affiliate-level panel data on the global operations of U.S.-based multinational firms from the Bureau of Economic Analysis (BEA) during 1989–2008. Our main empirical analysis evaluates firms in the computer industry (SIC 357).[4] Importantly, the data include separate direct measures of parent- and affiliate-specific R&D

[1]National Science Board (2014).

[2]U.S. Securities and Exchange Commission Form 10-K, Western Digital Corporation (2014).

[3]Intermediate inputs innovated and manufactured at Western Digital's U.S. sites include magnetic head wafers which, with further processing, are capable of reading from and recording onto data-storage media. See Igami (2015).

[4]We also study the pharmaceutical (SIC 283) and motor vehicles (SIC 371) industries and three broader associated sectors: industrial machinery (SIC 35), chemicals (SIC 28), and transportation equipment (SIC 37).

spending within each firm. The availability of panel data that include measures of production inputs, output, and R&D by site within the same firm for a series of years is both essential for answering our research question and also unusual.

Our estimates indicate that the performance of an affiliate is persistent and increasing in its own R&D investment, in line with results in Aw, Roberts, and Xu (2011), Doraszelski and Jaumandreu (2013) and Bøler, Moxnes, and Ulltveit-Moe (2015). 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 is consistent with Cohen and Levinthal (1989), which proposes that R&D enhances a firm's ability to assimilate and exploit information—including information resulting from R&D performed at the parent site, in this case. We also find that omitting parent innovation as a determinant of affiliate performance leads to an overstatement—of approximately 50 percent—in the performance impact of affiliates' own R&D. Our results further suggest that the flow of tangible goods from parents to affiliates is an important channel though which productivity gains resulting from R&D investments are transmitted to foreign sites, consistent with Keller and Yeaple (2013). By contrast, we do not find evidence that affiliate R&D determines performance growth at other sites within the firm.

Our estimates have quantitative implications for the measured value-added return to U.S. parent R&D investment as well as for the distribution of performance across production sites within a multinational firm. A direct implication is that the firm-level gains from parent innovation exceed the parent-level gains. Specifically, we find that the gross private return to parent 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 approximately 33 percent for the average firm. In addition, our estimates 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 58 percent; by contrast, eliminating an affiliate's own R&D would result in an analogous decline of just 7.8 percent. This impact of parent innovation on long-run affiliate performance further translates into an effect on affiliate value added, thus implying that countries hosting high levels of U.S. affiliate activity would observe a significant decline in GDP if affiliates were suddenly unable to benefit from U.S. parent innovation. These results suggest multinationals' parent innovation increases output growth abroad, an effect that may be further magnified by spillovers from affiliates to domestic firms (Javorcik 2004a).

A potential concern with our estimates is that multinationals may misreport R&D at the site level. The possibility of intentional overreporting of R&D spending among affiliates located in high-tax countries (and vice versa), could affect the consistency of our estimates. To account for this possibility, we use information on U.S. states' R&D tax credit policies from Wilson (2009) and foreign countries' intellectual property rights from Ginarte and Park (1997) and Park (2008) to build instruments for parent and affiliate R&D expenditures, and re-estimate the model parameters using these instruments. This alternative approach reveals that our baseline estimates are conservative.[5]

This paper contributes to a large literature evaluating the impact of R&D investment on plant-

[5]We also evaluate specifications that exclude affiliates located in tax havens, and we find results similar to our main estimates.

level or domestic firm-level outcomes; Hall, Mairesse, and Mohnen (2010) provide a detailed review. Methodologically, our analysis follows Aw, Roberts, and Xu (2011) and Doraszelski and Jaumandreu (2013) in that we model R&D investment as shifting firm performance. Our use of R&D policy incentives to identify the impact of innovation on performance further relates our estimation approach to Bøler, Moxnes, and Ulltveit-Moe (2015). We contribute to this research by evaluating the affiliate-level impact of headquarters innovation, and find that parent R&D is an important source of performance gain among affiliates within the same firm, particularly for those affiliates that also innovate and import from their U.S. parent.[6]

Our results are closely related to studies assessing the implications of input trade and proprietary technology transfer within multinational firms.[7] Specifically, our paper is relevant for work aiming to establish the existence of such transfer across plants within the multinational firm (Branstetter, Fisman, and Foley 2006; Keller and Yeaple 2013; Gumpert 2015).[8] A distinction relative to this literature is that we infer the flow of technology by estimating the impact of U.S. parent R&D investment on affiliate performance, and therefore do not rely on observed proxies for technology transfer. As a result, we are able to evaluate the information content of these proxies. Specifically, our results indicate that the payment of royalties and license fees is positively correlated with the affiliate-level impact of parent innovation, but we also find that affiliates reporting no such payments benefit from parent innovation on average.

Finally, 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; Tintelnot 2015). Our results indicate this distribution depends on endogenous decisions of the multinational firm, including the supply of inputs by parents to affiliates and the extent of parent and affiliate R&D investments. The estimates we recover, together with the levels of R&D spending and intrafirm trade observed in the data, are able to rationalize the finding that affiliates are significantly less productive than their parents (see Tintelnot 2015, and Head and Mayer 2015).

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 and 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. Section 8 concludes. Derivations and additional details may be found in the online Appendix.

[6]Our results thus suggest parent innovation may contribute to affiliates' productivity advantage relative to unaffiliated firms, relating our study to Doms and Jensen (1998) and Guadalupe, Kuzmina, and Thomas (2012).

[7]See, for example, Helpman (1984), Burstein and Monge-Naranjo (2009), McGrattan and Prescott (2010), McGrattan (2012), Irrazabal, Moxnes, and Opromolla (2013), Bilir (2014), and Ramondo, Rappaport, and Ruhl (2015).

[8]Branstetter, 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. For work providing indirect evidence of intrafirm trade in intangibles within U.S. multiplant firms, see Giraud (2013) and Atalay, Hortaçsu, and Syverson (2014).

2 Empirical Model

This section describes an empirical model of production and innovation investment in the multinational firm. Within the firm, each affiliate production site is associated with an idiosyncratic performance index that reflects both the productivity and demand shifter of the site.[9] In our baseline specification, performance at an affiliate site may be shifted separately by its own R&D investment and by that of its parent. Importantly, this model yields estimating equations that enable us to recover the parameters determining the impacts of parent and affiliate innovation on affiliate performance using available data.

2.1 Setup

Time is discrete. Consider a set of multinational firms $i = 1, \ldots, \mathcal{I}_t$ operating within the same manufacturing industry. 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 affiliates. We describe below the demand, production, and firm performance of foreign affiliates, and postpone the 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 any two distinct affiliates j and j' of firm i operate in distinct markets n_{ij} and $n_{ij'}$.[10] Importantly, the country corresponding to n_{ij} is the firm-i, affiliate-j production location, but need not be the location of its customers.[11]

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

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

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.[12] Market-level variables $P_{n_{ij}t}$ and $Q_{n_{ij}t}$ denote the period-t aggregate price index and aggregate demand level, respectively.

[9] Performance combines supply and demand shifters as in Foster, Haltiwanger and Syverson (2008), De Loecker (2011), Roberts et al (2011), and Bøler, Moxnes, and Ulltveit-Moe (2015).

[10] This definition is consistent with the data, in which all variables are observed at the firm-country-industry level. Note that this market definition accommodates a range of firm structures, from those with affiliates selling the same good in spatially segmented markets, to those with affiliates selling different goods in the same geographic market.

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

[12] Recent 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 by differences in σ. We extend the model in Appendix A.5 to allow σ to vary by year; estimates appear in section 6. Allowing for variable markups given the demand in (1) involves relaxing the assumed monopolistic competition structure; see Appendix A.6 for discussion.

2.3 Production

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

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

where

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

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

and $\boldsymbol{\alpha} = (\alpha_l, \alpha_k, \alpha_{ll}, \alpha_{kk}, \alpha_{lk})$.[13] 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. This production function combines materials with a translog function of capital and labor inputs, defined in (3) and (4), according to a Cobb-Douglas technology.[14] The elasticity of output with respect to materials is captured in (2) by α_m. Equation (4) allows output elasticities with respect to capital and labor to be heterogeneous across affiliates, 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 price is common to all affiliates within a market-year: $P_{ijt}^m = P_{n_{ij}t}^m$.[15] The materials use index M_{ijt} may include inputs sourced by an affiliate from its U.S. parent.[16]

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

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

where $\boldsymbol{\beta} = \boldsymbol{\alpha}(1-\alpha_m)\iota$ and $\iota = (\sigma - 1)/(\sigma - \alpha_m(\sigma - 1))$; see Appendix A.1 for a derivation of this expression. In (5), $\kappa_{n_{ij}t}$ is a market-year parameter that accounts for the materials price $P_{n_{ij}t}^m$, aggregate price index $P_{n_{ij}t}$, and aggregate demand level $Q_{n_{ij}t}$ 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 physical productivity and demand (product quality) shock: $\psi_{ijt} \equiv \iota(\omega_{ijt} + \xi_{ijt})$. We refer to ψ_{ijt}

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

[14] The 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 those in the literature (e.g. De Loecker and Warzynski 2012). See Appendix A.7 for a discussion of the empirical challenge of estimating a production function that is translog in unobserved materials.

[15] Labor and capital prices thus vary freely across affiliates. Assuming 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).

[16] Our assumption that $P_{n_{ij}t}^m$ varies at the market-year level would in this case be consistent with a model in which the cost of inputs sourced by an affiliate from its parent is the product of parent-country production costs and an iceberg trade cost that varies across affiliate countries. In section 5 below, we explore the importance of affiliate imports from the parent as potentially shaping the impact of parent innovation on affiliate performance.

as performance.

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

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

and assume that measurement error ε_{ijt} is mean independent of all variables known to firm i at t

$$\mathbb{E}_t(\varepsilon_{ijt}) = 0. \tag{7}$$

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

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

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

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

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 $\mu_{n_{ij}t}$ reflecting country-sector-year characteristics that affect the performance of affiliates operating in n_{ij} at t.[17] We define $r_{ijt} = 0$ for observations in which affiliate j does not perform R&D.[18] 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 marginal distribution of η_{ijt} beyond the mean independence implied by (8). In section 6, we allow η_{ijt} to be correlated across sites j within firm i to capture firm-specific shocks.

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).[19] A

[17] Equations (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).

[18] In 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 2). In section 6, we expand the specification in (9) by adding the interaction between $\mu_{n_{ij}t}$ and an indicator variable equal to 1 if affiliate j of firm i has positive R&D expenditure at $t-1$, and zero otherwise.

[19] In 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^2_{ijt-1} in (9); due to the resulting incidental parameters problem, this would cause asymptotic bias in our parameter estimates. See Appendix A.8 for details.

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; these include a generalization of (9) that permits affiliate imports from the parent to influence affiliate performance.

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 $t+1$, $\mathcal{J}_{it+1}$.[20] These decisions are a function of firm i's state vector $\mathbf{S}_{it}$, which has elements

$$S_{ijt} = (\psi_{ijt}, K_{ijt}, P^l_{ijt}, P^k_{ijt}, P^m_{n_{ij}t}, Q_{n_{ij}t}, P_{n_{ij}t}, \mu_{n_{ij}t}, \chi^k_{ijt}, \chi^r_{ijt}, F_{ijt}), \tag{10}$$

where χ^k_{ijt} and χ^r_{ijt} 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

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

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

$$\Pi(S_{ijt}, I_{ijt}, L_{ijt}, M_{ijt}, P_{ijt}, R_{ijt}) = VA^*_{ijt} - W^l_{ijt} - C_k(P^k_{ijt}, I_{ijt}, K_{ijt}, \chi^k_{ijt}) - C_r(R_{ijt}, \chi^r_{ijt}) - F_{ijt} \tag{12}$$

where $C_k(\cdot)$ and $C_r(\cdot)$ are cost functions of investment in physical capital and R&D, and $W^l_{ijt} \equiv P^l_{ijt} L_{ijt}$ 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^k_{ijt}, χ^r_{ijt}, and F_{ijt}; we thus follow Doraszelski and Jaumandreu (2013) by leaving these functions unspecified.[21]

[20] The baseline model thus assumes affiliate entry and exit decisions are taken with a one-period lag; section 4.3 considers instead the case of instantaneous affiliate entry and exit. For any X, we henceforth denote $\mathbf{X}_{it} = \{X_{ijt}\}_{j \in \mathcal{J}_{it}}$.

[21] The absence of restrictions on $C_k(\cdot)$, $C_r(\cdot)$ and the distributions of χ^k_{ijt} and χ^r_{ijt} ensures that there exist $C_k(\cdot)$ and $C_r(\cdot)$ functions as well as χ^k_{ijt} and χ^r_{ijt} 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. This poses a well-known computational challenge (e.g. Holmes 2011, Morales, Sheu, and Zahler 2015).

3 Data and Measurement

Estimating the parameters of the model described in section 2 requires measures of R&D investment for each firm parent, and measures of inputs, value added, and R&D expenditures for each of its affiliates.

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 data provide information on U.S. parent companies and each foreign affiliate on an annual basis.[22] 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.2.

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^l_{ijt}. 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. The index of materials use M_{ijt}, the output market price and quantity indexes $P_{n_{ij}t}$ and $Q_{n_{ij}t}$, and the materials price $P^m_{n_{ij}t}$ are not observed. Importantly, the data include separate parent- and affiliate-level measures of R&D investment for distinct sites within the same firm.[23] The availability of panel data that include measures of production inputs, value added, and R&D by site for a series of years 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}.[24] This measure does not account for spending on capital inputs, routine testing and quality control, market research, or legal expenses related to patents.[25]

Labor L_{ijt} in the model is a production input, but a plant performing innovation may dedicate

[22]The 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.

[23]R&D investment is recorded for all majority-owned affiliates that are above a minimum size (Appendix A.2).

[24]Unfortunately, this precludes separately estimating the impact of product and process innovation; see Cohen and Klepper (1996) and Dhingra (2013) for models that feature both types of innovation.

[25]In benchmark years, the data further 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, when available, 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).

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 6.2.

We estimate the parameters of the model separately by industry. Multinationals are assigned to an industry based on that of the parent. For each firm, our main analysis uses only information on manufacturing 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 6.5, 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 1. Section 6.6 presents estimates for multinational firms in Motor Vehicles and Equipment (SIC 371) and Pharmaceutical Drugs (SIC 283). To assess the sensitivity of our results to industrial differences across sample firms, we also present estimates in section 6 that rely on a broader industry definition: Chemicals (SIC 28), Industrial Machinery (SIC 35), and Transportation Equipment (SIC 37).[26]

3.2 Descriptive Statistics

Table 2 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: first, most U.S. parents invest in R&D; second, only 47.8 percent of these 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. In addition, affiliates tend to account for only a small share—8.5 percent, on average—of firm-level R&D spending. Parent sites are thus responsible for the substantial majority of U.S.-based multinational firms' investment in innovation. Furthermore, R&D spending is more concentrated at the parent site than production; as Table 2 shows, the affiliate share in total firm sales and employment is three times larger than the affiliate share in total firm R&D investment, and the affiliate share in total firm value added is twice as large as the affiliate share in total firm R&D investment.

4 Estimation

This section derives estimating equations using the model in section 2. We show how these equations may be combined with the data described in section 3 to estimate the parameters of the model.

[26]These are the three largest manufacturing sectors by number of affiliates (U.S. BEA 2008).

4.1 Estimation Approach

The parameters of interest are those determining affiliate value added in (6), $\boldsymbol{\beta}$, and those governing the evolution of 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

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

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 $\gamma_{n_{ij}t} \equiv \mu_{n_{ij}t} + \kappa_{n_{ij}t} - \rho\kappa_{n_{ij}t-1}$ is a market-year effect that accounts for both the unobserved quantity and prices embedded in $\kappa_{n_{ij}t}$ and the market-year specific component of firm performance, $\mu_{n_{ij}t}$.

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}, giving rise to 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 associated with (11) and (12) 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 2013). Conditional on these first-stage 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 $\{\gamma_{n_{ij}t}\}$. 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 is

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

where, as in (12), W^l_{ijt} 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

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

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 $(\beta_l, \beta_{ll}, \beta_{lk})$.[27] With the estimates

[27]Note that (14) is compatible with differences in value added per worker across affiliates. These may be due to either differences in wages P^l_{ijt} 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 locations even within the same country-sector pair n. And, because we do not restrict the correlation between affiliate performance ψ_{ijt} and wages P^l_{ijt}, our model is consistent with differential labor market frictions across affiliates depending on performance ψ_{ijt} (e.g. due

$(\hat{\beta}_l, \hat{\beta}_{ll}, \hat{\beta}_{lk})$ in hand, we recover an estimate of the measurement error ε_{ijt} for each firm i, affiliate j, and period t: $\hat{\varepsilon}_{ijt} = va_{ijt} - w^l_{ijt} + \log(\hat{\beta}_l + 2\hat{\beta}_{ll} l_{ijt} + \hat{\beta}_{lk} k_{ijt})$.[28]

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

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

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

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

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 $\{\gamma_{n_{ij}t}\}$ using the Frisch-Waugh-Lovell theorem.

This second step identifies the impact of affiliate and parent R&D investment on affiliate performance by exploiting variation in r_{ijt-1} and r_{i0t-1} conditional on lagged affiliate performance ψ_{ijt-1} and market-year unobserved effects $\gamma_{n_{ij}t}$. 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 may arise if, for example, performance is persistent and r_{ijt-1} is determined by ψ_{ijt-1}. In addition, market-year fixed effects $\gamma_{n_{ij}t}$ account for country characteristics that may affect both affiliate performance 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. 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 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 $t-1$.

Our approach thus allows us to estimate the impact of R&D by comparing performance growth across 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

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 performance ψ_{ijt}, the capital price P^k_{ijt}, or capital adjustment costs χ^k_{ijt}.

[28] Theoretically, an alternative way of handling measurement error in value added would be to use a control function for unobserved performance ψ_{ijt} using either capital (Olley and Pakes 1996) or labor (Levinsohn and Petrin 2003). 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 information set; this is infeasible in our setting due to the large dimensionality of $\mathbf{S}_{it}$ in (10). Using labor as a control variable would require assuming affiliate wages P^l_{ijt} are mean independent of affiliate performance ψ_{ijt}; it has been shown that this restriction impacts production function estimates substantially (Collard-Wexler and De Loecker 2015, and De Loecker at al 2015). Our approach has the advantage of not relying on such a restriction.

performance and a market-year, note that optimal R&D investments of the affiliate and its parent may differ due to differences in their realized R&D cost shocks χ^r_{ijt-1} and χ^r_{i0t-1} (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.

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.[29]

4.2 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.[30]

Differences between actual and reported value added are accommodated by the model through the term ε_{ijt}. 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 section 6.3 below, we assess the robustness of our main results to patterns of value added misreporting that do not verify (15). Specifically, we re-estimate the model using two subsamples. First, we exclude affiliates in countries with exceptionally low effective corporate tax rates (Gravelle 2015). All else equal, if misreporting of value added for tax purposes is present in our data, the set of excluded affiliates in 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 excluded affiliates in 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). If the parameter estimates found using these two subsamples are similar, it would suggest that such value added misreporting is unlikely to affect our results.

Regarding R&D spending, the model in section 2 presumes actual and reported R&D expenditures coincide. Suppose instead 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})

[29] Consistency 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 assumption 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 $t-1$.

[30] See Hines and Rice (1994), Hines (1997), and Bernard, Jensen, and Schott (2006).

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.3).[31] To obtain consistent estimates in the presence of misreporting, we re-estimate the model parameters replacing the estimator described in 4.1 with a Generalized Method of Moments (GMM) estimator that relies on policy instruments for identification.

Our first policy instrument combines information on innovation incentives faced by U.S. parent sites and on 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 $\text{IPR}_{n_{ij}t}$ in the affiliate host country at t.[32] The relevance of this instrument is confirmed in the data. As the estimates in Table A.1 (Appendix A.3) show, a) an increase in the user cost of R&D in a U.S. state reduces R&D spending by multinational parents located in that state, and raises R&D spending by their foreign affiliates to an extent increasing in local intellectual property rights; and b) an increase in intellectual property rights in a foreign country has a larger positive impact on R&D spending by local affiliates when their U.S. parent is located in a state with a relatively high user cost of R&D.

The validity of this instrument relies on its mean independence of other policies that influence R&D misreporting x_{ijt-1} by affiliates, after controlling for the covariates (other than r_{ijt-1} and r_{i0t-1}) in the conditioning set in (17). Importantly, any unobserved 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 across countries, as it would be controlled for by the set of country-sector-year fixed effects $\{\gamma_{n_{ij}t}\}$ included in (16). Notice this is true even if the impact of such an omitted variable on the incentives to misreport varies across sectors.

Our second policy instrument is the measured local user cost of R&D, UCRD_{it}. Similar to Bloom, Shankerman and Van Reenen (2013), there is a strong negative correlation between U.S. parent R&D investment and UCRD_{it} in our sample. For our instrument to be valid, UCRD_{it} must also be uncorrelated with the extent of parent R&D misreporting x_{i0t}. Claiming that x_{i0t} and UCRD_{it} are mean-independent requires a stronger assumption than in the case of $\text{UCRD}_{it} \times \text{IPR}_{n_{ij}t}$: 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 in this case cause downward bias in our estimates of the affiliate performance elasticity with respect to parent R&D spending (see Appendix A.3 for additional details).

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

[31] A 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.

[32] We obtain UCRD_{it} from Wilson (2009), and $\text{IPR}_{n_{ij}t}$ from Ginarte and Park (1997) and Park (2008). For details on the construction of the variable $\text{IPR}_{n_{ij}t}$, see Appendix A.2. This index is widely used; see, for example, Javorcik (2004b), McCalman (2004), Branstetter, Fisman, and Foley (2006), Qian (2007), and Bilir (2014).

conservative estimates of the elasticity with respect to parent R&D.[33]

4.3 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 takes place. 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 for 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 the definition of η_{ijt} in (8) implies entry events prior to t are independent of η_{ijt}.

By contrast, 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.[34] 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).[35]

5 Main Results

5.1 Baseline Model

Estimates corresponding to the baseline model in section 2 appear in Table 3. Results in columns 1 through 4 are estimated following the approach in section 4.1; columns 5 through 8 present optimal two-step GMM estimates that use the instruments discussed in section 4.2. Details on these two estimators are in Appendix A.9.

The estimates in column 1 reveal that the performance of an affiliate is both persistent and increasing in its own R&D spending, confirming that affiliates of U.S. multinational firms are in these ways qualitatively similar to firms studied in the previous literature (Aw, Roberts and Xu 2011, Doraszleski and Jaumandreu 2013). Column 2 adds innovation by the U.S. parent of each affiliate, accommodating the possibility that proprietary technology developed by the parent impacts

[33] Endogeneity could also arise if affiliate R&D is correlated with other (unobserved) investments in future performance; for example, affiliates may acquire technology from external sources. Our instrumentation strategy addresses this form of endogeneity provided that, within the set of affiliates in a market-year, these other technology investments are independent over time and not correlated with the instrument $\text{UCRD}_{it} \times \text{IPR}_{n_{ij}t}$.

[34] By contrast, notice that 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.

[35] Theoretically, 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 $\mathbf{S}_{it}$.

affiliates within the same firm. The estimates indicate parent innovation is a highly significant determinant of affiliate-level performance: specifically, all else equal, the impact of a one-percent increase in parent R&D on affiliate performance is approximately four times larger than the impact of a one-percent increase in affiliate R&D spending.[36] Also, while the estimated coefficient on affiliate R&D investment remains highly significant in column 2, notice that its magnitude declines when we account for parent innovation. Omitting parent innovation, as in column 1, thus leads to an overstatement—of approximately 50 percent—in the performance impact of affiliates' own R&D. Column 3 also considers the possible importance of innovation by other foreign affiliates within the same firm. The estimates suggest the strong centrality of parent innovation within the firm network: parent innovation affects the performance of all firm affiliates but, by contrast, the impact of innovation by an affiliate is limited to its own site.

Column 4 shows that parent and affiliate R&D investment are complementary. 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. This complementarity implies affiliates do not benefit equally from parent innovation.

The estimates in columns 1 through 4 are thus consistent with a view of the multinational firm featuring a central innovating parent that affects all of its affiliates (some more than others), and peripheral affiliates whose innovation does not affect other sites within the same firm. In what follows, we therefore emphasize specifications that include both own-affiliate and parent R&D spending, as in Table 3, columns 2 and 4.

Columns 5 through 8 of Table 3 evaluate these two specifications using instruments that account for possible R&D investment misreporting (see section 4.2 for a description of these instruments). The estimates in columns 5 and 6 suggest that those in columns 2 and 4 are conservative. As discussed in section 4.2, this pattern is consistent with parents and affiliates over-reporting R&D in locations with stronger R&D incentives. The difference between the NLS and GMM estimates is larger for the coefficient on affiliate R&D than for that corresponding to parent R&D, also consistent with the hypotheses developed in section 4.2 and Appendix A.3. Columns 7 and 8 further indicate that the GMM estimates are largely robust to the introduction of U.S. state fixed effects corresponding to the parent location: the positive coefficient on the interaction between parent and affiliate R&D is the only estimate sensitive to including these additional controls.

Finally, the lower rows of Table 3 present the average value added elasticities with respect to labor and capital, and their respective standard deviations across affiliates. Importantly, our model predicts that these elasticities will be lower than the corresponding output elasticities for

[36] Notice that distinct parent and affiliate R&D coefficients (μ_a and μ_p) are compatible with optimizing firms that equalize net R&D returns across sites. Specifically, under the assumptions in section 2, our estimates imply decreasing returns to R&D performed at each site. Suppose that the cost of performing R&D were identical across sites within a firm. The larger estimated coefficient on parent R&D in column 2 of Table 3 ($\hat{\mu}_p > \hat{\mu}_a$) would thus indicate that, to be consistent with optimality, parent sites should perform more R&D than affiliates. As described in section 3, U.S. parents do indeed invest in significantly more R&D than their foreign affiliates. The R&D cost shifters χ^r_{ijt} included in (10) also affect the distribution of optimal R&D investment across the sites of a multinational firm by allowing for flexible (fixed and variable) cost differences between parent and affiliate sites.

each affiliate.[37] The estimates reported in Table 3 are thus compatible with affiliate production functions exhibiting approximately constant returns to scale.[38]

5.2 Intrafirm Trade

Parents may share proprietary technology with affiliates either by embedding it in tangible goods exported to affiliates or by communicating it in intangible forms (Keller and Yeaple, 2013; Irrazabal, Moxnes, and Opromolla, 2013). Inputs imported from a parent 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 complementarity between affiliates' intrafirm imports and R&D investment, we extend the specification in (9) to account for observed affiliate imports from their U.S. parents.

Specifically, Table 4 presents estimates of a variant of the model described in section 2 in which (9) is replaced by

$$\mathbb{E}_{t-1}[\psi_{ijt}] = \rho\psi_{ijt-1} + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \mu_{pm} r_{i0t-1} im_{ijt-1} + \mu_{apm} r_{ijt-1} r_{i0t-1} im_{ijt-1} + \mu_m im_{ijt-1} + \mu_{n_{ij}t}, \tag{18}$$

where im_{ijt-1} is the log of imports received by firm i's affiliate j from its U.S. parent at $t-1$. 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. 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 4 than in specifications that do not control for im_{ijt-1} (column 2, Table 3), consistent with complementarity and thus in line with Bøler, Moxnes, and Ulltveit-Moe (2015).

Columns 3 and 4 of Table 4 evaluate the hypothesis in Keller and Yeaple (2013) and Irrazabal, Moxnes, and Opromolla (2013) 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 in fact shared primarily 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

[37] As $\sigma > 1$ and $0 < \alpha_m < 1$, the parameters of the value added function $\boldsymbol{\beta}$ in (5) are always smaller in absolute value than the corresponding production function parameters $\boldsymbol{\alpha}$.

[38] Specifically, for an elasticity of substitution σ around 5 (e.g. Head and Mayer 2015) and a materials expenditure share in sales α_m around 0.65 (Doraszleski 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.

innovation, even in the absence of imported inputs.

The estimates in column 4 imply a distribution across affiliates in the elasticity of performance with respect to parent R&D spending that appears in Figure 1. This figure reveals substantial heterogeneity across affiliates in the impact of parent R&D, reflecting heterogeneity in affiliates' own R&D spending and in the volume of inputs they import from their U.S. parents. At the lower end of the distribution, for affiliates that neither innovate nor import inputs from their U.S. parent, parent R&D has an insignificant impact. However, as affiliates increase either their own R&D investment or import larger input volumes from their parent, this elasticity rises sharply. Figure 1 also plots the elasticity of affiliate performance with respect to affiliates' own R&D investment, which is substantially smaller than the estimated elasticity with respect to parent R&D. Comparing the two distributions, the importance of parent R&D in determining the long-run performance of multinational affiliates is apparent. Further evaluation and implications appear in section 7.1.

6 Alternative Specifications

6.1 Intrafirm Technology Licensing

One interpretation of observed technology royalties and license fees flowing within and across firms is that they are an exact proxy for otherwise unobserved technology transfer (Hines 1995, Branstetter, Fisman, and Foley 2006). Table 5 presents estimates on the extent to which royalties paid by an affiliate to its U.S. parent reflect the affiliate-level impact of R&D performed by the parent.

The estimates in Table 5 indicate that, while royalties paid by affiliates to parents are correlated with the impact of parent R&D on affiliate performance, these payments do not fully capture this impact. Specifically, estimated coefficients on royalty payments in columns 1 and 2 are positive and statistically significant, consistent with the idea that these payments reflect affiliate acquisition of technology developed by the U.S. parent that results in performance gains by the affiliate. However, Table 5 further indicates that R&D performed by the U.S. parent also has a positive and statistically significant impact on affiliates that do not report paying royalties or license fees to their parent firm, indicating an effect of parent innovation not captured by observed royalty payments.

6.2 Labor Measurement

As discussed in section 3, concerns may arise regarding the measure of labor inputs in production. Estimates in Table 3 measure labor 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 new measure of production workers. Estimates using this new measure appear in columns 1 and 2 of Table 6, and reveal estimates similar to the baseline results in Table 3.[39]

[39] Appendix A.2 contains additional details regarding the construction of our measure of production workers.

6.3 Value Added Measurement

Because firms may misreport value added in response to tax incentives (section 4.2), we consider two robustness checks that assess the potential importance of such misreporting. These additional tests exclude affiliates that are more likely to either overstate or understate value added, all else equal. Specifically, firms may have the incentive to overstate value added earned by affiliates in countries with exceptionally low corporate tax rates. To assess the potential importance of this for our estimates, we exclude affiliates located in countries identified as tax havens in Gravelle (2015).[40] The resulting estimates appear in columns 3 and 4 in Table 6 and are very similar to those in Table 3. One mode through which affiliates may understate value added is by overstating spending on inputs purchased within the firm (transfer pricing). To assess the potential relevance of this for our results, we restrict the sample of affiliates to include only those with an intrafirm input share near zero; all else equal, such affiliates are less likely to understate value added. The estimates appear in columns 5 and 6 in Table 6; the results are qualitatively unchanged. Taken together, the estimates from these two subsamples suggest that the results in section 5 are unlikely to be driven by systematic value added misreporting.

6.4 Alternative Performance Process

The stochastic process for affiliate performance ψ_{ijt} in (9) may be restrictive in ways that affect the estimates described in section 5; here, we assess the sensitivity of these estimates to several aspects of this specification.

First, equation (9) implicitly assumes that the direct impact of innovation on affiliate performance occurs one year after the R&D investment date. However, certain technology development projects may require additional time. 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 7 and are qualitatively similar to those in Table 3. The main estimates are thus not sensitive to this adjustment in the assumed timing of innovation impact.

Second, section 3 indicates that, while almost all U.S. parents perform R&D, only the minority of foreign affiliates do so (24 percent). We therefore estimate an alternative model that relaxes the assumption in (9) that all affiliates, even those choosing the corner solution of zero R&D spending, are affected equally by the market-year factors accounted for by the performance process effects $\{\mu_{n_{ij}t}\}$. Specifically, we extend (9) to include an additional set of effects $\{\tilde{\mu}_{n_{ij}t} \times d_{ijt-1}\}$ where d_{ijt-1} is an indicator variable that is equal to one if firm i's affiliate j performs R&D during $t-1$. This allows for flexible differences in average performance growth between innovating and non-innovating affiliates located within the same market and year. The resulting estimates are in columns 5 to 8 of Table 7; we obtain nearly identical results to those reported in section 5.

Third, the performance process in (9) relies on 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

[40]See Appendix A.2 for the list of countries identified as tax havens in Gravelle (2015).

affiliate performance terms in periods prior to $t-1$; these may be correlated with measures of R&D investment included as covariates in (16). Specifically, serial correlation in affiliate R&D implies omitting relevant lags of affiliate performance causes upward bias in estimates of 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 8. While the coefficient on affiliate R&D weakens, those on parent R&D and its interaction with affiliate R&D remain essentially unchanged.

Finally, a multinational firm i may experience shocks to productivity or product quality that simultaneously affect the performance of all firm-i affiliate sites. To accommodate this possibility, columns 5 and 6 of Table 8 provide estimates that allow performance shocks η_{ijt} to be correlated across affiliates within the same multinational firm and year. The results are qualitatively identical to those in columns 2 and 4 of Table 3.

6.5 Heterogeneous Innovation Impact

The results in section 5 are based on a model in which all affiliates share identical parameters governing production, demand, and performance evolution. However, the impact of parent R&D on affiliate performance may depend on the industrial proximity of parent and affiliate. While the estimates in Table 3 use information on all manufacturing affiliates of sample firms, columns 1 and 2 in Table 9 restrict the sample to include only those affiliates operating within the computer industry (SIC 357), the same sector as the U.S. parent. The estimates reveal that restricting the composition of affiliates in this way does not affect the average impact of parent R&D, but does significantly increase the complementarity between parent and affiliate innovation.

Appendix A.5 shows that heterogeneity in the demand elasticity σ implies the affiliate performance effects of parent and affiliate R&D are heterogeneous. In columns 3 through 6 of Table 9, we allow the demand elasticity parameter σ to vary by year with the aim of accounting for changes in the substitutability of products manufactured by computer firms that occur during the sample period. We obtain percentiles of the distribution of the affiliate performance elasticity with respect to R&D spending by multiplying the estimates in columns 3 through 6 of Table 9 with percentiles from the distribution of ι_{2t} reported in the last column. Comparing the estimates in columns 1 through 4 in Table 3 with the elasticity distributions implied by the estimates in Table 9, we find that the estimates computed under the assumption of homogeneous σ are very close to the averages of the corresponding elasticities computed allowing σ to vary by year.[41]

6.6 Other Industries

Table 10 assesses the relevance of the results discussed above for affiliates 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

[41] To characterize the distribution across years for each elasticity, multiply the corresponding estimate in columns 3 through 6 by the quantiles of the distribution of ι_2 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 additional details.

(SCI 371). One important difference is that, in the motor vehicles sector, parent and affiliate R&D are not complementary. To determine whether our results are sensitive to the industry definition, we estimate the model parameters within each of the three associated broader sectors: industrial machinery (SIC 35), chemicals (SIC 28), and transportation equipment (SIC 37).[42] Table 10 Panel B shows estimates corresponding to machinery and chemicals firms, and indicate that main qualitative results do not depend on the degree of aggregation in the estimation sample.

7 Quantitative Implications

The parameters estimated in sections 5 and 6 have implications for the impact of site-specific R&D investments within a multinational firm on the long-run performance of its production affiliates. These parameters also mediate the contribution of headquarters innovation to the overall value added earned by a multinational firm, as well as the aggregate value added of foreign countries that host firm affiliates. Throughout this section, we rely on the estimates of (18) appearing in column 4, Table 4, in which affiliate performance responds to parent R&D, affiliate R&D, and affiliate imports from the parent.

7.1 Headquarters Innovation and Affiliate Performance

To assess the impact of parent innovation on long-run affiliate performance, we define the expected, long-run performance of firm i's affiliate j as $\psi_{ij} \equiv \mathbb{E}\big[\lim_{s\to\infty} \psi_{ijs} \big| r_{i0t}, r_{ijt},\, im_{ijt}\big]$, where we assume for simplicity that firm i expects r_{i0t}, r_{ijt}, and im_{ijt} to remain constant at their respective period-t values for a specified base year t. In the absence of parent innovation, holding fixed all other performance determinants, this expected long-run performance becomes $\psi_{ij,r_0} \equiv \mathbb{E}\big[\lim_{s\to\infty} \psi_{ijs} \big| 0, r_{ijt}, im_{ijt}\big]$.[43] The long-run, relative affiliate-j performance impact of removing the contribution of its parent R&D is then $\Delta_{ij,r_0} \equiv \psi_{ij,r_0} - \psi_{ij}$. We define analogous respective impacts of eliminating affiliate R&D Δ_{ij,r_j} and intrafirm imports Δ_{ij,im_j} (Appendix A.10).

Figure 2 plots the distributions of Δ_{ij,r_0}, Δ_{ij,r_j}, and Δ_{ij,im_j} across affiliates in the base year $t = 2004$.[44] At each percentile, these three distributions are ordered: $\Delta_{ij,r_0} < \Delta_{ij,im_j} < \Delta_{ij,r_j}$; long-run affiliate performance is thus lowest in the absence of parent innovation.

These distributions also reveal the economic significance of parent innovation for firm affiliates: for the median affiliate, eliminating the impact of parent R&D would, all else equal, 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. For the median affiliate, headquarters innovation thus appears to be seven times as important as affiliate R&D and five times as important as imports from the parent in determining its long-run performance. However, the relative impacts of eliminating these three sources of

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

[43] Accounting for the impact of a change in parent R&D on optimal affiliate R&D and imports from the parent would require fully specifying the dynamic problem of firm i and solving the optimization in (11).

[44] Appendix A.10 shows that Δ_{ij,r_0}, Δ_{ij,r_j}, and Δ_{ij,im_j} may be evaluated using data observed at t and the estimates in section 5.2. 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.

performance gain also differ across affiliates. Specifically, Figure 2 reveals a decline in long-run affiliate performance of more than 80 percent for the quintile of affiliates most impacted by losing access to parent R&D. Comparable declines in performance are 25 percent for affiliate access to imports from the parent and 20 percent for affiliates' own R&D investment.

7.2 Innovation and the Headquarters Performance Advantage

Recent work including Tintelnot (2015) and Head and Mayer (2015) has concluded that affiliate productivity is systematically below that of corresponding parent sites. Using the empirical framework in section 2 and estimates from section 5.2, we assess the specific contribution of R&D investment and its differential impact across firm sites to the gap in parent-affiliate performance.

Evaluating the expected long-run performance of each U.S. parent requires first estimating the determinants of parent performance evolution. To do this, we assume parents face demand and production functions similar to (1) through (4), but governed by distinct parameter values σ and $\boldsymbol{\alpha}$. We also assume parent performance evolves according to (8) with

$$\mathbb{E}_{t-1}[\psi_{i0t}] = \rho_0 \psi_{i0t-1} + \mu_0 r_{i0t-1} + \mu_{a0} r_{i-0t-1} + \mu_{n_{i0}t}, \tag{19}$$

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 and GMM estimates of ρ_0, μ_0, and μ_{a0} appear in Table A.2.[45] These estimates show that parent R&D is an important determinant of its own performance, while R&D performed by firm-i foreign affiliates is not a significant determinant of parent performance.[46]

Notice that, in the model, the relative long-run performance difference $\psi_{ij} - \psi_{i0}$ between the parent of firm i and its affiliate j depends on the respective sequences of market-year specific performance shocks affecting these two sites, $\{\mu_{n_{i0}s}\}$ and $\{\mu_{n_{ij}s}\}$. However, given the data described in section 3, these unobserved effects cannot be separately identified from the market-year unobserved terms $\{\kappa_{n_{i0}s}\}$ and $\{\kappa_{n_{ij}s}\}$ in the parent and affiliate value added functions. We therefore focus here on the parent-affiliate performance gap determined by the impact of parent R&D, affiliate R&D, and affiliate imports from the parent. Appendix A.11 shows that this component may be evaluated using affiliate-level estimates from section 5.2, parent-level estimates from Table A.2, and data for a base year t. The resulting distribution of the difference $\psi_{ij} - \psi_{i0}$ appears in Figure 3 for the base year $t = 2004$. Values are uniformly below 0.50, indicating all affiliates exhibit lower long-run performance levels due to innovation and intrafirm trade than their respective parent sites.[47] Fur-

[45] Note that the GMM estimates of μ_0 are larger than their corresponding NLS estimates, consistent with the discussion in section 4.2.

[46] Table A.2 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.

[47] To build intuition, consider a simpler setting with $\mu_a = \mu_{ap} = \mu_m = \mu_{apm} = 0$. Comparing the estimates in Tables 3 and A.2, 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 $\hat{\mu}_p/(1 - \hat{\rho}) = 0.0164/(1 - 0.778) = 0.74$. Thus, if parent R&D were the only determinant of performance, parents

thermore, the median value is 0.33 and the distribution is compressed, ranging between 0.27 and 0.37 for approximately 60 percent of affiliates. Our estimates thus indicate that multinationals' choices regarding site-specific R&D spending and intrafirm input supply generate large, endogenous performance disadvantages for affiliates relative to parents.

7.3 Gross Return to Parent R&D

The model described in section 2 assumes that parents of U.S. multinationals determine R&D investment optimally. Provided $R_{i0t} > 0$, the optimal R&D investment of the parent of firm i satisfies the following condition

$$\frac{\partial V(\mathbf{S}_{it})}{\partial R_{i0t}} = \underbrace{\frac{\partial \sum_{s>t} \sum_{j \in \mathcal{J}_{is}} \delta^{s-t} \mathbb{E}_t[VA^*_{ijs}]}{\partial R_{i0t}}}_{\equiv GR_{i0t}} - \underbrace{\frac{\partial C_r(R_{i0t}, \chi^r_{i0t})}{\partial R_{i0t}}}_{\equiv MC_{i0t}} = 0, \tag{20}$$

where GR_{i0t} is the gross firm-i return to parent R&D and MC_{i0t} is its marginal cost.[48] The optimal level of headquarters R&D spending by firm i thus depends on the shape of the gross return function GR_{i0t}; moreover, firms i 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 5.2 together imply that, at any 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 volume of trade flows from the firm-i parent to each of its affiliates at t, c) the R&D expenditure of firm-i foreign affiliates at t, d) expected growth at t in the variables mentioned in a) through c). We quantify the combined contributions of a), b), and c) to the gross, firm-i return to parent R&D, assuming for simplicity that firm i expects the number of affiliates and affiliate value added, R&D spending, and imports from the parent to remain stable at period-t levels.[49] In this case,

$$GR_{i0t} = \frac{\partial \mathbb{E}_t[\psi_{i0t+1}]/\partial r_{i0t}}{1-\rho_0} \frac{VA^*_{i0t}}{R_{i0t}} + \sum_{\substack{j \in \mathcal{J}_{it} \\ j \neq 0}} \frac{\partial \mathbb{E}_t[\psi_{ijt+1}]/\partial r_{i0t}}{1-\rho} \frac{VA^*_{ijt}}{R_{i0t}}, \tag{21}$$

where $\partial \mathbb{E}_t[\psi_{i0t+1}]/\partial r_{i0t} = \mu_0$ and $\partial \mathbb{E}_t[\psi_{ijt+1}]/\partial r_{i0t} = \mu_p + \mu_{ap} r_{ijt} + \mu_{pm} im_{ijt} + \mu_{apm} r_{ijt} im_{ijt}$.

The distribution of GR_{i0t} across firms i appears in Figure 4 for the base year $t = 2004$. The average firm-wide gross return is 1.8 dollars and the median is 0.48 dollars.[50] The share of the

would be 75% more productive than their affiliates in the long run. The difference between this number and the distribution in Figure 3 reveals the influence of affiliate innovation and intrafirm trade on the performance gap.

[48] Parent R&D is almost always positive in our estimation sample (see section 3). As in Doraszelski and Jaumandreu (2013), we account for the possibility that R_{i0t} may not capture all costs associated with R&D performance. This is consistent with the data, as our measure of R&D spending excludes capital innovation costs.

[49] Detailed derivations appear in Appendix A.12.

[50] Notice that we have not imposed the equality in (20) when computing the estimates in sections 5 and 6. However, the estimates in section 5.2 indicate that the gross long-run return is near one for most firms, implying that the effective marginal cost of investing one additional dollar in R&D is also close to one for most firms.

gross return to firm-i parent innovation that may be attributed to its affiliates is

$$\Lambda_{it} \equiv 1 - \Big(\frac{\mu_0}{1-\rho_0}\frac{VA^*_{i0t}}{R_{i0t}}\Big/GR_{i0t}\Big),$$

and the distribution of Λ_{it} appears in Figure 5 for the base year $t = 2004$. The average value of Λ_{it} across firms i is 23 percent, though it varies widely. This affiliate share is just 3 percent for the 25th-percentile firm, rising to 38 percent at the 75th percentile, and up to 61 percent at the 90th percentile. For comparison, the distribution of the affiliate share in total firm value added is 5 percent for the 25th-percentile firm, 31 percent at the 75th percentile, and just 47 percent at the 90th percentile.[51] The affiliate share in firm value added is thus less dispersed than the affiliate share in value added gains from parent R&D; the latter also exceeds the former in the majority of firms. The distinction between these distributions reflects, in part, the influence of affiliate R&D investment and imports from the parent, which amplify value added gains from parent R&D.[52]

Accounting for the impact of U.S. parent R&D on foreign affiliates within the same firm therefore has a substantial impact on the measured gross return to parent R&D. This result also implies, through (20), that multinational firms with more extensive affiliate operations optimally choose higher levels of parent R&D investment, as do firms with innovative, import-intensive affiliates.

That affiliates influence optimal parent innovation decisions has further implications for the responsiveness of parent innovation to U.S. R&D policy incentives. For example, to stimulate private innovation investment, U.S. state and federal governments subsidize innovation through R&D tax credits (Wilson 2009). Such subsidies are captured by the model in section 2 through the parent R&D cost shifter χ^r_{i0t}, which impacts optimal firm-i parent R&D: equation (20) implies this impact hinges on the gross return function (21). Specifically, totally differentiating (20) with respect to χ^r_{i0t} and noting that $\partial GR_{i0t}/\partial\chi^r_{i0t} = 0$ yields

$$\frac{dR_{i0t}}{d\chi^r_{i0t}} = \frac{\partial MC_{i0t}}{\partial\chi^r_{i0t}}\Bigg/\Bigg(\frac{\partial GR_{i0t}}{\partial R_{i0t}} - \frac{\partial MC_{i0t}}{\partial R_{i0t}}\Bigg). \tag{22}$$

Second order conditions imply that the denominator in (22) is negative.[53] The impact of a policy change in χ^r_{i0t} on optimal parent innovation by firm i is thus 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

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

[52]Intuitively, note that our model implies the distributions of Λ_{it} and Λ^{va}_{it} coincide if parent and affiliate performance are equally persistent ($\rho_0 = \rho$), and share the same elasticity with respect to parent R&D ($\mu_p = \mu_0$, $\mu_{ap} = \mu_{pm} = \mu_{apm} = 0$). In our estimates, these restrictions are not upheld, and it is thus not surprising that the distributions are distinct. To build intuition for our findings, maintain the assumption that $\mu_{ap} = \mu_{pm} = \mu_{apm} = 0$, but use the estimates $\hat{\mu}_p$, $\hat{\mu}_0$, $\hat{\rho}_0$, and $\hat{\rho}$ appearing in Tables 3 and A.2. In this case,

$$\Lambda_{it} = \frac{\frac{\mu_p}{1-\rho}\sum_{j\in\mathcal{J}_{it},j\neq 0} VA^*_{ijt}}{\frac{\mu_0}{1-\rho_0}VA^*_{i0t} + \frac{\mu_p}{1-\rho}\sum_{j\in\mathcal{J}_{it},j\neq 0} VA^*_{ijt}} = \frac{\frac{\mu_p}{1-\rho}\Lambda^{va}_{it}}{\frac{\mu_0}{1-\rho_0}(1-\Lambda^{va}_{it}) + \frac{\mu_p}{1-\rho}\Lambda^{va}_{it}} = \frac{0.7387\Lambda^{va}_{it}}{1.2917(1-\Lambda^{va}_{it}) + 0.7387\Lambda^{va}_{it}}.$$

This expression implies that $\Lambda_{it} < \Lambda^{va}_{it}$. However, our estimates of μ_{ap}, μ_{pm}, and μ_{apm} further imply a higher long-run affiliate performance elasticity with respect to parent R&D than obtains when $\mu_{ap} = \mu_{pm} = \mu_{apm} = 0$: thus, for firms with pervasive affiliate innovation and imports from the parent, Λ_{it} exceeds Λ^{va}_{it}.

[53]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. governments' R&D subsidies increasing the volume of R&D by parents, as found in Table A.1 (Appendix A.3).

influence of parent innovation on affiliate sites abroad, accounting for this influence directly affects the predicted efficacy of U.S. innovation policy (see Appendix A.12 for additional details).

7.4 U.S. Parent Innovation and GDP Growth Abroad

The affiliate-level estimates in section 5 indicate parent innovation, through its positive impact on affiliate performance, raises the value added earned by foreign affiliates within the same multinational firm. An implication of this result 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 temporarily prevents U.S. firm affiliates from receiving the intrafirm benefit of parent innovation. 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.[54] 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 21 percent for the average foreign country.[55] For the average country, this decline is equal to 1.46 percent of its aggregate 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 and imports from the parent. Countries located near the United States experience larger effects: GDP in the computer industry would decline in Canada by 8.11 percent and in Mexico by 2.26 percent. The impact is also more pronounced for countries with a computer industry dominated by U.S. affiliates, including Finland (10.54 percent decline) and Indonesia (2.40 percent decline).

8 Conclusion

This paper evaluates the impact of R&D investment on affiliate performance within the multinational firm. Using data on the global activity of U.S.-based firms and a dynamic model of multinational innovation, our estimates indicate that U.S. parent R&D investment systematically increases foreign affiliate performance. Parent innovation is the primary determinant of long-run affiliate performance, and that, for the average U.S.-based firm, approximately one quarter of the overall gross return to parent R&D accrues to foreign affiliates of the same firm. However, affiliates do not benefit equally: the impact of parent R&D on affiliate performance increases systematically in affiliates' own R&D spending, and in the volume of inputs they import from the parent. As a consequence, it is not only the size of a multinational firm that determines its incentive to pursue R&D: instead, we find that the joint distribution of value added, intrafirm trade, and R&D spending across sites mediates the impact of headquarters R&D spending on overall firm performance.

In the United States, multinational parents account for the majority of firm R&D spending,

[54] Specifically, we consider the impact before a) affiliates adjust labor, investment in physical capital, and R&D spending, b) active foreign domestic firms adjust output, c) entry by either foreign domestic firms or affiliates of non-U.S. multinationals occurs. Assessing the total impact of such a policy would also require accounting for possible spillover effects of changes in U.S. affiliate performance on foreign domestic plants (e.g. Javorcik 2004a).

[55] See Appendix A.13 for details.

and our estimates thus reveal a channel through which policies aimed at stimulating U.S. R&D investment also raise firm performance in developing countries hosting U.S. firm affiliates. This cross-border effect may be larger than implied by our estimates if knowledge and productivity spillovers from U.S. firms' foreign affiliates to domestic firms are important, as found in Javorcik (2004a) and Poole (2013). Importantly, if the benefits resulting from increased affiliate productivity are not fully captured by U.S. entities, our estimates point to a specific mechanism through which U.S. innovation policy may result in a cross-border externality. Such gains from U.S. innovation are, nevertheless, more pronounced for countries with extensive and innovative U.S. affiliate operations, suggesting the importance of further work evaluating the policy determinants and impacts of multinational activity from the perspective of developing countries. For this, our results point to potential productivity growth advantages among countries that simultaneously implement R&D stimulus and import openness policies as foreign direct investment liberalization proceeds.

REFERENCES

[1] **Adams, James D., and Adam B. Jaffe.** (1986). "Bounding the Effects of R&D: An Investigation Using Matched Establishment-Firm Data." *RAND Journal of Economics*, vol. 27, no. 5, pp. 700–721.

[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] **Atalay, Enghin, Ali Hortaçsu, and Chad Syverson.** (2014). "Vertical Integration and Input Flows." *American Economic Review*, vol. 104, no. 4, pp. 1120–1148.

[4] **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.

[5] **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.

[6] **Bernard, Andrew, J., Bradford Jensen, and Peter Schott.** (2006). "Transfer Pricing by U.S.-Based Multinational Firms." Yale University mimeo.

[7] **Bilir, L. Kamran.** (2014). "Patent Laws, Product Lifecycle Lengths, and Multinational Activity." *American Economic Review*, vol. 104, no. 7, pp. 1979–2013.

[8] **Bloom, Nicholas, Mark Schankerman, and John Van Reenen.** (2013). "Identifying Technology Spillovers and Product Market Rivalry." *Econometrica*, vol. 81, no. 4, pp. 1347–1393.

[9] **Bøler, Esther A.** (2015). "Technology-skill Complementarity in a Globalized World." University of Oslo mimeo.

[10] **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.

[11] **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.

[12] **Caves, Richard.** (2007). *Multinational Enterprise and Economic Analysis*, Cambridge University Press.

[13] **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.

[14] **Cohen, Wesley M., and David A. Levinthal.** (1989). "Innovation and Learning." *Economic Journal*, vol. 99, pp. 569–596.

[15] **Collard-Wexler, Allan and Jan De Loecker.** (2005). "Reallocation and Technology: Evidence from the U.S. Steel Industry" *American Economic Review*, vol. 105, no. 1, pp. 131–171.

[16] **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.

[17] **De Loecker, Jan, and Frederic Warzynski.** (2012). "Markups and Firm Level Export Status." *American Economic Review*, vol. 102, no. 6, pp. 2437–2471.

[18] **De Loecker, Jan.** (2013). "Detecting Learning by Exporting." *American Economic Journal: Microeconomics*, vol. 5, no. 3, pp. 1–21.

[19] **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.

[20] **Dharmapal, Dhammika, and James Hines Jr.** (2009). "Which Countries Become Tax Havens?" *Journal of Public Economics*, vol. 93, no. 9-10, pp. 1058–1068.

[21] **Dhingra, Swati.** (2013). "Trading Away Wide Brands for Cheap Brands." *American Economic Review*, vol. 103, no. 6, pp. 2554–2584.

[22] **Doms, Mark E., and J. Bradford Jensen.** (1998). "Comparing Wages, Skills, and Productivity between Domestically and Foreign-Owned Manufacturing Establishments in the United States," in Baldwin, R., Lipsey, R., Richards, D. (editors), *Geography and Ownership as Bases for Economic Accounting*, University of Chicago Press.

[23] **Doraszelski, Ulrich, and Jordi Jaumandreu.** (2013). "R&D and Productivity: Estimating Endogenous Productivity." *Review of Economic Studies*, vol. 80, pp. 1338–1383.

[24] **Doraszelski, Ulrich, and Jordi Jaumandreu.** (2015). "Measuring the Bias of Technological Change." University of Pennsylvania mimeo.

[25] **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.

[26] **Gandhi, Amit, Salvador Navarro, and David Rivers.** (2013). "On the Identification of Production Functions: How Heterogeneous is Productivity?" University of Wisconsin-Madison mimeo.

[27] **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.

[28] **Giraud, Xavier.** (2013). "Proximity and Investment: Evidence from Plant-level Data." *The Quarterly Journal of Economics*, vol. 128, no. 2, pp. 861-915.

[29] **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.

[30] **Gravelle, Jane G.** (2015). "Tax Havens: International Tax Avoidance and Tax Evasion." *Congressional Research Service Report*, no. 7-5700, pp. 605-627.

[31] **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.

[32] **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.

[33] **Guadalupe, María, Olga Kuzmina, and Catherine Thomas.** (2012). "Innovation and Foreign Ownership." *American Economic Review*, vol. 102, no. 7, pp. 3594–3627.

[34] **Gumpert, Anna.** (2015). "The Organization of Knowledge in Multinational Firms." University of Munich mimeo.

[35] **Hall, Robert E., and Dale Jorgenson.** (1967). "Tax Policy and Investment Behavior." *American Economic Review*, vol. 57, no. 3, pp. 391–414.

[36] **Hall, Bronwyn H., Jacques Mairesse, and Pierre Mohnen.** (2010). "Measuring the Returns to R&D." University of California–Berkeley mimeo.

[37] **Head, Keith, and Thierry Mayer.** (2015). "Brands in Motion: How Frictions Shape Multinational Production." University of British Columbia mimeo.

[38] **Helpman, Elhanan.** (1984). "A Simple Theory of International Trade with Multinational Corporations." *Journal of Political Economy*, vol. 92, no. 3, pp. 451–471.

[39] **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.

[40] **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.

[41] **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.

[42] **Holmes, Thomas J.** (2011). "The Diffusion of Wal-Mart and Economies of Density." *Econometrica*, vol. 79, n. 1, pp. 253–302.

[43] **Igami, Mitsuru.** (2015). "Estimating the Innovator's Dilemma: Structural Analysis of Creative Destruction in the Hard Disk Drive Industry." *Journal of Political Economy*, forthcoming.

[44] **Irarrazabal, 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.

[45] **Jaumandreu, Jordi, and Heng Yin.** (2014). "Cost and Product Advantages: A Firm-level Model for Chinese Exports and Industry Growth." Boston University mimeo.

[46] **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.

[47] **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.

[48] **Keller, Wolfgang, and Stephen R. Yeaple.** (2013). "The Gravity of Knowledge." *American Economic Review*, vol. 103, no. 4, pp. 1414–1444.

[49] **Levinsohn, James, and Amil Petrin.** (2003). "Estimating Production Functions Using Inputs to Control for Unobservables." *Review of Economic Studies*, vol. 70, pp. 317–341.

[50] **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.

[51] **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.

[52] **McGrattan, Ellen R.** (2012). "Transition to FDI Openness: Reconciling Theory and Evidence." *Review of Economic Dynamics*, vol. 15, no. 4, pp. 437–458.

[53] **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.

[54] **Morales, Eduardo, Gloria Sheu, and Andres Zahler.** (2015). "Extended Gravity." Princeton University mimeo.

[55] **National Science Board.** (2014). *Science and Engineering Indicators 2014.* Arlington, VA: National Science Foundation (NSB 14-01).

[56] **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.

[57] **Olley, Steven G., and Ariel Pakes.** (1996). "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, vol. 64, n. 6, pp. 1263–1295.

[58] **Park, Walter G.** (2008). "International Patent Protection: 1960–2005," *Research Policy*, vol. 37, no. 4, pp. 761–766.

[59] **Poole, Jennifer P.** (2013). "Knowledge Transfers from Multinational to Domestic Firms: Evidence from Worker Mobility," *Review of Economics and Statistics*, vol. 95, no. 2, pp. 393–406.

[60] **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.

[61] **Ramondo, Natalia, Veronica Rappaport, and Kim Ruhl.** (2015). "Intra-firm Trade and Vertical Fragmentation in U.S. Multinational Corporations." *Journal of International Economics, forthcoming.*

[62] **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.

[63] **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.

[64] **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.

[65] **Tintelnot, Felix.** (2015). "Global Production with Export Platforms." University of Chicago mimeo.

[66] **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.

[67] **Western Digital Corporation.** (2014). 10-K Report. Retrieved from http://www.sec.gov/Archives/edgar/data/106040/000119312512361018/d356205d10k.htm.

[68] **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.

Table 1: Summary Statistics, 1989-2008

	Parent-Year Level		Affiliate-Year Level	
Variable	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 2: Descriptive Statistics, R&D Allocation in the Multinational Firm

	Firm-Year Level	
Variable	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 3: Baseline Estimates

	NLS				GMM			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Persistence	0.7915^a	0.7721^a	0.7722^a	0.7681^a	0.6749^a	0.7221^a	0.7109^a	0.6763^a
	(0.0095)	(0.0098)	(0.0098)	(0.0098)	(0.0327)	(0.0354)	(0.0267)	(0.0338)
Affiliate R&D	0.0064^a	0.0042^a	0.0043^a	-0.0272^a	0.0365^a	-0.2704^b	0.0293^a	0.2493^b
	(0.0013)	(0.0013)	(0.0014)	(0.0070)	(0.0122)	(0.1140)	(0.0110)	(0.1215)
Parent R&D		0.0164^a	0.0166^a	0.0103^a	0.0326^b	-0.0791^b	0.0322^c	0.0890^a
		(0.0022)	(0.0023)	(0.0026)	(0.0161)	(0.0377)	(0.0169)	(0.0345)
Other Affiliates' R&D			-0.0003					
			(0.0013)					
Affiliate R&D × Parent R&D				0.0025^a		0.0240^a		-0.0160
				(0.0005)		(0.0087)		(0.0086)
Labor Elasticity	0.4682	0.4682	0.4682	0.4682	0.4682	0.4682	0.4682	0.4682
	(0.0646)	(0.0646)	(0.0646)	(0.0646)	(0.0646)	(0.0646)	(0.0646)	(0.0646)
Capital Elasticity	0.2526	0.2488	0.2487	0.2447	0.2245	0.2094	0.2115	0.2041
	(0.0572)	(0.0536)	(0.0536)	(0.0517)	(0.0545)	(0.0437)	(0.0436)	(0.0463)
Observations	4,194	4,194	4,194	4,194	4,194	4,194	4,194	4,194

Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. Columns (1) to (4) report Nonlinear Least Squares estimates of the parameters in (9) and several variants thereof; columns (5) to (8) 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 parenthesis. 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 $\beta_l + \beta_{ll} 2 l_{ijt} + \beta_{lk} k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk} 2 k_{ijt} + \beta_{lk} l_{ijt}$. 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.

Table 4: Estimates with Affiliate Imports from U.S. Parent

	(1)	(2)	(3)	(4)
Persistence	0.7601^{a}	0.7568^{a}	0.7532^{a}	0.7520^{a}
	(0.0099)	(0.0099)	(0.0099)	(0.0099)
Affiliate R&D	0.0029^{b}	-0.0248^{a}	0.0026^{b}	-0.0146^{c}
	(0.0013)	(0.0070)	(0.0013)	(0.0077)
Parent R&D	0.0122^{a}	0.0070^{a}	0.0020	0.0009
	(0.0023)	(0.0026)	(0.0030)	(0.0033)
Affiliate R&D × Parent R&D		0.0022^{a}		0.0013^{c}
		(0.0005)		(0.0007)
Parent Imports	0.0085^{a}	0.0082^{a}	-0.0185^{a}	-0.0140^{a}
	(0.0012)	(0.0012)	(0.0055)	(0.0057)
Parent R&D × Parent Imports			0.0022^{a}	0.0018^{a}
			(0.0004)	(0.0005)
Affiliate R&D × Parent R&D × Parent Imports				0.0000
				(0.0000)
Labor Elasticity	0.4682	0.4682	0.4682	0.4682
	(0.0646)	(0.0646)	(0.0646)	(0.0646)
Capital Elasticity	0.2393	0.2360	0.2300	0.2290
	(0.0513)	(0.0498)	(0.0482)	(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 (18) 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 $\beta_l + \beta_{ll} 2 l_{ijt} + \beta_{lk} k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk} 2 k_{ijt} + \beta_{lk} l_{ijt}$. 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 5: Estimates with Affiliate Royalty Payments to U.S. Parent

	(1)	(2)	(3)	(4)
Persistence	0.7463^{a}	0.7445^{a}	0.7411^{a}	0.7398^{a}
	(0.0101)	(0.0101)	(0.0102)	(0.0102)
Affiliate R&D	0.0027^{b}	-0.0203^{a}	0.0024^{c}	-0.0173^{b}
	(0.0013)	(0.0060)	(0.0013)	(0.0073)
Parent R&D	0.0128^{a}	0.0085^{a}	0.0072^{a}	0.0038
	(0.0022)	(0.0025)	(0.0026)	(0.0029)
Affiliate R&D × Parent R&D		0.0018^{a}		0.0016^{a}
		(0.0005)		(0.0006)
Parent Royalties	0.0112^{a}	0.0107^{a}	-0.0146^{a}	-0.0129^{c}
	(0.0012)	(0.0012)	(0.0062)	(0.0064)
Parent R&D × Parent Royalties			0.0020^{a}	0.0019^{a}
			(0.0005)	(0.0005)
Affiliate R&D × Parent R&D × Parent Royalties				-0.0000
				(0.0000)
Labor Elasticity	0.4682	0.4682	0.4682	0.4682
	(0.0646)	(0.0646)	(0.0646)	(0.0646)
Capital Elasticity	0.2370	0.2350	0.2345	0.2295
	(0.0499)	(0.0488)	(0.0487)	(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 $\beta_l + \beta_{ll} 2 l_{ijt} + \beta_{lk} k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk} 2 k_{ijt} + \beta_{lk} l_{ijt}$. 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 6: Labor and Value Added Measurement

	Non-R&D Labor		*No Tax Havens*		*Low Imports from Parent*	
	(1)	(2)	(3)	(4)	(5)	(6)
Persistence	0.7739^a	0.7702^a	0.7648^a	0.7639^a	0.8036^a	0.8002^a
	(0.0101)	(0.0101)	(0.0106)	(0.0106)	(0.0122)	(0.0122)
Affiliate R&D	0.0043^a	-0.0251^a	0.0056^a	-0.0163^b	0.0046^b	-0.0247^a
	(0.0013)	(0.0068)	(0.0014)	(0.0076)	(0.0018)	(0.0087)
Parent R&D	0.0136^a	0.0081^a	0.0182^a	0.0139^a	0.0128^a	0.0069^a
	(0.0021)	(0.0025)	(0.0024)	(0.0028)	(0.0027)	(0.0032)
Affiliate R&D × Parent R&D		0.0023^a		0.0017^a		0.0024^a
		(0.0005)		(0.0006)		(0.0007)
Labor Elasticity	0.4709	0.4709	0.4813	0.4813	0.4525	0.4525
	(0.0649)	(0.0649)	(0.0678)	(0.0678)	(0.0722)	(0.0722)
Capital Elasticity	0.2620	0.2585	0.2324	0.2284	0.1956	0.1900
	(0.0565)	(0.0550)	(0.0537)	(0.0514)	(0.0361)	(0.0337)
Observations	4,194	4,194	3,432	3,432	2,212	2,212

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) replace L_{ijt} with a measure of production employment; columns (3) and (4) exclude affiliates located in tax havens identified in Gravelle (2015); columns (5) and (6) include only affiliates with zero or low imports from their U.S. parent. 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 $\beta_l + \beta_{ll} 2 l_{ijt} + \beta_{lk} k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk} 2 k_{ijt} + \beta_{lk} l_{ijt}$. 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 7: 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]	0.8067[a]	0.8253[a]	0.8230[a]	0.7880[a]	0.7682[a]	0.7682[a]	0.7398[a]
	(0.0106)	(0.0106)	(0.0114)	(0.0114)	(0.0094)	(0.0097)	(0.0097)	(0.0097)
Affiliate R&D ($t-1$)					0.0230[a]	0.0189[a]	0.0190[a]	-0.0127
					(0.0034)	(0.0034)	(0.0034)	(0.0091)
Parent R&D ($t-1$)						0.0170[a]	0.0171[a]	0.0117[a]
						(0.0022)	(0.0024)	(0.0026)
Other Affiliates' R&D ($t-1$)							-0.0002	
							(0.0013)	
Affiliate R&D × Parent R&D ($t-1$)								0.0022[a]
								(0.0006)
Affiliate R&D ($t-2$)	0.0052[a]	-0.0239[a]						
	(0.0014)	(0.0081)						
Parent R&D ($t-2$)	0.0154[a]	0.0084[a]						
	(0.0027)	(0.0033)						
Affiliate R&D × Parent R&D ($t-2$)		0.0023[a]						
		(0.0006)						
Affiliate R&D ($t-3$)			0.0051[a]	-0.0140				
			(0.0015)	(0.0089)				
Parent R&D ($t-3$)			0.0162[a]	0.0114[a]				
			(0.0030)	(0.0037)				
Affiliate R&D × Parent R&D ($t-3$)				0.0015[b]				
				(0.0007)				
Labor Elasticity	0.4684	0.4684	0.4706	0.4706	0.4682	0.4682	0.4682	0.4682
	(0.0625)	(0.0625)	(0.0613)	(0.0613)	(0.0646)	(0.0646)	(0.0646)	(0.0646)
Capital Elasticity	0.2427	0.2397	0.2311	0.2297	0.2381	0.2344	0.2344	0.2334
	(0.0506)	(0.0496)	(0.0473)	(0.0466)	(0.0508)	(0.0478)	(0.0478)	(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 parenthesis. Persistence corresponds to estimates of ρ. Labor Elasticity is the average value of $\beta_l + \beta_{ll} 2 l_{ijt} + \beta_{lk} k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk} 2 k_{ijt} + \beta_{lk} l_{ijt}$. 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: Additional Performance Lags and Correlated Performance Shocks

	Additional Productivity Lags				*Firm Shocks*	
	(1)	(2)	(3)	(4)	(5)	(6)
Persistence ($t-1$)	0.7206^a	0.7162^a	0.6535^a	0.6497^a	0.7721^a	0.7681^a
	(0.0155)	(0.0155)	(0.0222)	(0.0222)	(0.0255)	(0.0255)
Persistence ($t-2$)	0.1267^a	0.1256^a	0.1475^a	0.1443^a		
	(0.0151)	(0.0151)	(0.0267)	(0.0266)		
Persistence ($t-3$)			0.0881^a	0.0882^a		
			(0.0211)	(0.0211)		
Affiliate R&D	0.0013	-0.0255^a	0.0010	-0.0237^a	0.0042^b	-0.0272^a
	(0.0013)	(0.0070)	(0.0016)	(0.0090)	(0.0017)	(0.0097)
Parent R&D	0.0145^a	0.0094^a	0.0136^a	0.0098^a	0.0164^a	0.0103^b
	(0.0022)	(0.0026)	(0.0027)	(0.0031)	(0.0039)	(0.0051)
Affiliate R&D × Parent R&D		0.0021^a		0.0019^a		0.0025^a
		(0.0005)		(0.0007)		(0.0008)
Labor Elasticity	0.4689	0.4689	0.4761	0.4761	0.4682	0.4682
	(0.0620)	(0.0620)	(0.0589)	(0.0589)	(0.0646)	(0.0646)
Capital Elasticity	0.2390	0.2364	0.2503	0.2487	0.2488	0.2447
	(0.0407)	(0.0402)	(0.0383)	(0.0384)	(0.0536)	(0.0517)
Observations	3,442	3,442	1,984	1,984	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 parenthesis. Persistence corresponds to estimates of ρ. Labor Elasticity is the average value of $\beta_l + \beta_{ll} 2 l_{ijt} + \beta_{lk} k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk} 2 k_{ijt} + \beta_{lk} l_{ijt}$. 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 9: Industry Restriction and Markup Heterogeneity

	Parent Sector		Heterogeneous Markups				*Distribution of* ι_{2t}	
	(1)	(2)	(3)	(4)	(5)	(6)		
Persistence	0.8271^a	0.8266^a	0.7982^a	0.7808^a	0.7808^a	0.7833^a	minimum	2.07
	(0.0253)	(0.0247)	(0.0089)	(0.0092)	(0.0093)	(0.0093)		
Affiliate R&D	0.0017	-0.1015^a	0.0021^a	0.0015^a	0.0015^a	-0.0026^c	percentile 25	2.63
	(0.0046)	(0.0261)	(0.0004)	(0.0004)	(0.0004)	(0.0014)		
Parent R&D	0.0166^a	-0.0002		0.0049^a	0.0049^a	0.0041^a	median	2.84
	(0.0075)	(0.0085)		(0.0007)	(0.0000)	(0.0008)		
Other Affiliates' R&D					0.0000		percentile 75	3.33
					(0.0004)			
Affiliate R&D × Parent R&D		0.0077^a				0.0003^a	maximum	3.72
		(0.0019)				(0.0001)		
Labor Elasticity	0.4249	0.4249	0.4702	0.4702	0.4702	0.4702	average	2.91
	(0.0935)	(0.0935)	(0.0831)	(0.0831)	(0.0831)	(0.0831)		
Capital Elasticity	0.1437	0.1212	0.1986	0.1974	0.1975	0.1961	std. dev.	0.40
	(0.1167)	(0.1058)	(0.0503)	(0.0488)	(0.0488)	(0.0483)		
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 ι_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 $\beta_l + \beta_{ll} 2 l_{ijt} + \beta_{lk} k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk} 2 k_{ijt} + \beta_{lk} l_{ijt}$. 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 10: Other Sectors

Panel A: Other 3-digit SIC Sectors

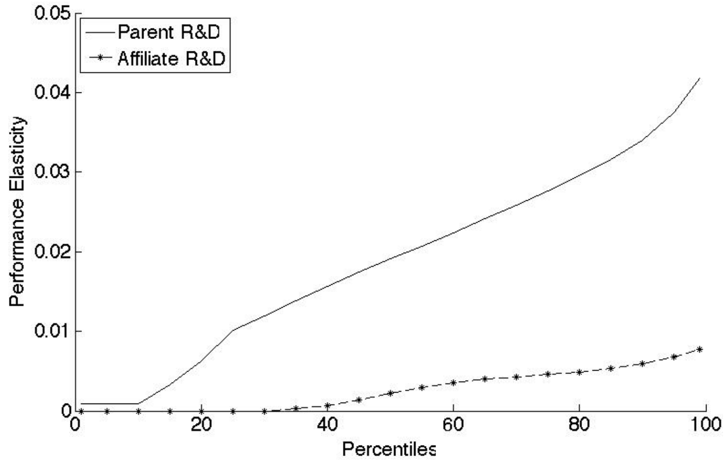
	Drugs				*Motor Vehicles and Motor Vehicle Equipment*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Persistence	0.7910^a	0.7879^a	0.7879^a	0.7860^a	0.6755^a	0.6696^a	0.6693^a	0.6696^a
	(0.0072)	(0.0073)	(0.0073)	(0.0073)	(0.0067)	(0.0068)	(0.0068)	(0.0068)
Affiliate R&D	0.0112^a	0.0108^a	0.0107^a	-0.0029	0.0035^a	0.0028^a	0.0026^a	-0.0004
	(0.0009)	(0.0009)	(0.0009)	(0.0064)	(0.0008)	(0.0008)	(0.0008)	(0.0038)
Parent R&D		0.0066^a	0.0055^a	0.0026^a		0.0087^a	0.0078^a	0.0081^a
		(0.0020)	(0.0023)	(0.0027)		(0.0012)	(0.0014)	(0.0014)
Other Affiliates' R&D			0.0011				0.0012	
			(0.0013)				(0.0009)	
Affiliate R&D × Parent R&D				0.0010^a				0.0003
				(0.0005)				(0.0003)
Labor Elasticity	0.4298	0.4298	0.4298	0.4298	0.5698	0.5698	0.5698	0.5698
	(0.0453)	(0.0453)	(0.0453)	(0.0453)	(0.0808)	(0.0808)	(0.0808)	(0.0808)
Capital Elasticity	0.1782	0.1779	0.1779	0.1778	0.2129	0.2092	0.2093	0.2084
	(0.0392)	(0.0386)	(0.0386)	(0.0247)	(0.0673)	(0.0641)	(0.0641)	(0.0637)
Observations	7,285	7,285	7,285	7,285	9,953	9,953	9,953	9,953

Panel B: 2-digit SIC Sectors

	Industrial and Commercial Machinery & Computer Equipment				*Chemicals and Allied Products*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Persistence	0.7188^a	0.7120^a	0.7106^a	0.7113^a	0.7469^a	0.7436^a	0.7427^a	0.7372^a
	(0.0092)	(0.0093)	(0.0094)	(0.0093)	(0.0060)	(0.0061)	(0.0061)	(0.0061)
Affiliate R&D	0.0049^a	0.0049^a	0.0048^a	-0.0053	0.0102^a	0.0096^a	0.0092^a	-0.0133^a
	(0.0011)	(0.0011)	(0.0011)	(0.0045)	(0.0007)	(0.0007)	(0.0007)	(0.0031)
Parent R&D		0.0101^a	0.0092^a	0.0064^a		0.0040^a	0.0021^c	-0.0005
		(0.0019)	(0.0020)	(0.0024)		(0.0010)	(0.0012)	(0.0012)
Other Affiliates' R&D			0.0013				0.0027^a	
			(0.0011)				(0.0008)	
Affiliate R&D × Parent R&D				0.0009^a				0.0019^a
				(0.0003)				(0.0002)
Labor Elasticity	0.5441	0.5441	0.5441	0.5441	0.4421	0.4421	0.4421	0.4421
	(0.0779)	(0.0779)	(0.0779)	(0.0779)	(0.0534)	(0.0534)	(0.0534)	(0.0534)
Capital Elasticity	0.1729	0.1682	0.1688	0.1674	0.1611	0.1604	0.1604	0.1609
	(0.0319)	(0.0283)	(0.0283)	(0.0279)	(0.0306)	(0.0301)	(0.0298)	(0.0293)
Observations	5,016	5,016	5,016	5,016	10,681	10,681	10,681	10,681

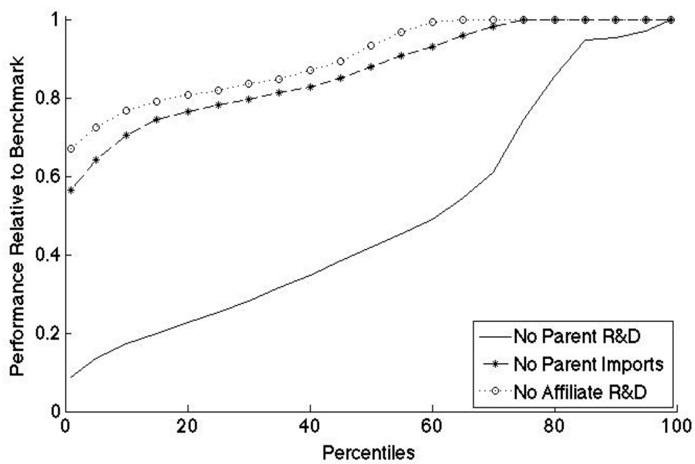
Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. All columns report Nonlinear Least Squares estimates corresponding to (16) as specified in section 2 and several variants thereof for firms in additional industries. In Panel A, columns (1) to (4) report estimates 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 $\beta_l + \beta_{ll} 2 l_{ijt} + \beta_{lk} k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk} 2 k_{ijt} + \beta_{lk} l_{ijt}$. 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.

Figure 1: Elasticities of Affiliate Performance



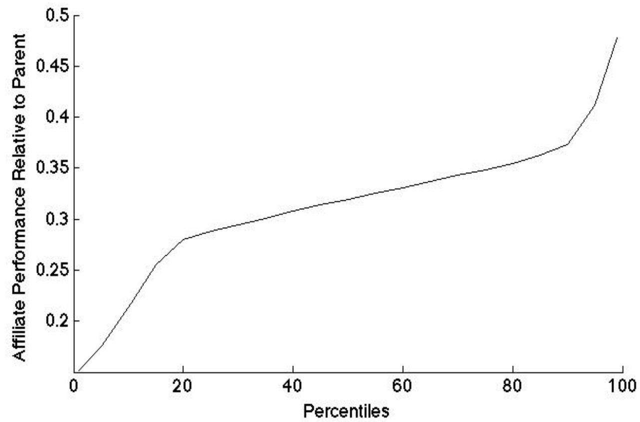
Notes: For each percentile indicated on the horizontal axis, the height of the solid line is the elasticity of affiliate performance with respect to parent R&D; and the height of the dashed line marked with asterisks indicates the elasticity of affiliate performance with respect to affiliate R&D.

Figure 2: Determinants of Long-Run Affiliate Performance Distribution



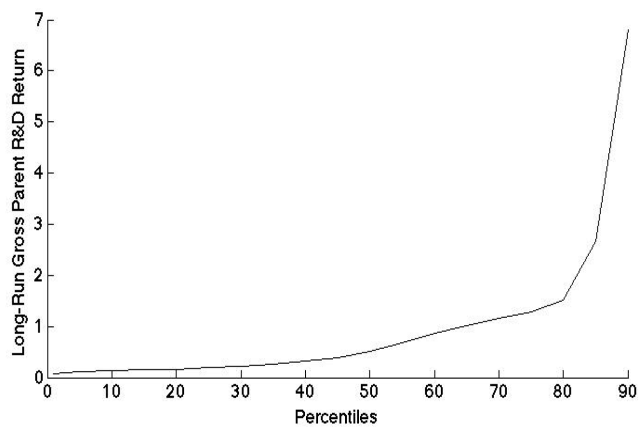
Notes: For each percentile indicated on the horizontal axis, the height of the solid line is the long-run affiliate performance without parent R&D relative to the long-run affiliate performance level in our benchmark specification; the height of the dashed line marked with asterisks indicates long-run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

Figure 3: Distribution of Affiliate Performance Relative to Firm Parent



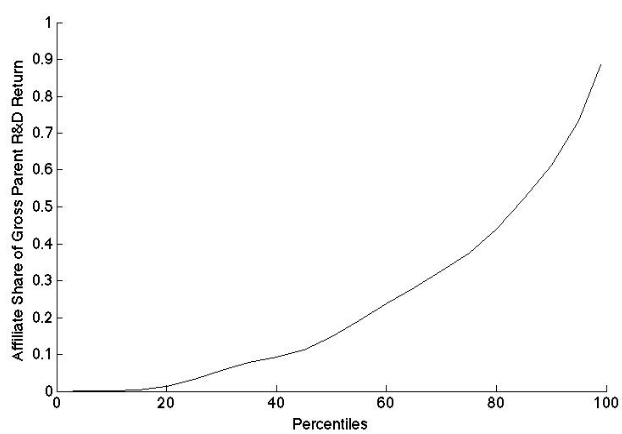
Notes: For each percentile indicated on the horizontal axis, the height of the solid line is the expected long-run affiliate performance due to affiliate and parent R&D spending and affiliate imports from the parent relative to the expected long-run parent performance due to parent R&D spending.

Figure 4: 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 5: 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.

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

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

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

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

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

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

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

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

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

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

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

$$P^m_{n_{ij}t}M^*_{ijt} = \left(\frac{\alpha_m(\sigma-1)}{\sigma}\right)Y_{ijt}, \tag{A.1}$$

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),

$$va^*_{ijt} = \kappa_{n_{ij}t} + h(k_{ijt}, l_{ijt}; \boldsymbol{\beta}) + \psi_{ijt},$$

where

$$\kappa_{n_{ij}t} = \ln\Big(1 - \frac{\alpha_m(\sigma - 1)}{\sigma}\Big) + \widetilde{\kappa}_{n_{ij}t}.$$

A.2 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).[56]

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.

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.

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), the U.S. Government Accountability Office (GAO), and Dharmapal and Hines (2009). 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

[56] During 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.

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.3 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 4.2, we define

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

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

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

where the error term is

$$\nu_{ijt} = \eta_{ijt} - \mu_a x_{ijt-1} - \mu_p x_{i0t-1} - \mu_{ap} x_{ijt-1} x_{i0t-1}. \tag{A.3}$$

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

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

As, by definition, $var_c(x_{ijt-1}) > 0$ and $var_c(r^*_{ijt-1} + x_{ijt-1}) > 0$, it is clear that $cov_c(x_{ijt-1}, r^*_{ijt-1}) \geq 0$ 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. $cov_c(x_{ijt-1}, subsidies_{ijt-1}) \geq 0$. 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. $cov_c(r^*_{ijt-1}, subsidies_{ijt-1}) \geq 0$. 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 $\text{UCRD}_{it} \times \text{IPR}_{n_{ij}t}$ 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

$$\frac{cov_c(\widehat{va}_{ijt}, \text{UCRD}_{it-1} \times \text{IPR}_{n_{ij}t-1})}{cov_c(r_{ijt-1}, \text{UCRD}_{it-1} \times \text{IPR}_{n_{ij}t-1})},$$

or, equivalently,

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

where, as above, $cov_c(\cdot)$ denotes the covariance after conditioning on the vector $(k_{ijt}, \widehat{va}_{ijt-1}, k_{ijt-1})$ 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 4.2 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

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

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

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

As long as the sign of $cov_c(x_{i0t-1}, \text{UCRD}_{it-1})$ is the same as that of $cov_c(r_{i0t-1}, \text{UCRD}_{it-1})$, we expect our GMM estimate of μ_p to be asymptotically downward biased. Specifically, as discussed in Section 4.2, 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 $cov_c(x_{i0t-1}, \text{UCRD}_{it-1}) < 0$ and $cov_c(r_{i0t-1}, \text{UCRD}_{it-1}) < 0$. 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 $\text{UCRD}_{it-1} \times \text{IPR}_{n_{ij}t-1}$ 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 and 2 in Table A.1 show that UCRD_{it} and $\text{UCRD}_{it} \times \text{IPR}_{n_{ij}t}$ have an impact on 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 (column 2). 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 (column 1). 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.[57]

Table A.1: Determinants of Parent and Affiliate R&D

	Correlations	
Dependent variable:	r_{ijt}	r_{i0t}
	(1)	(2)
$\text{UCRD}_{it} \times \text{IPR}_{n_{ij}t}$	0.741^a	-0.035
	(0.055)	(0.026)
UCRD_{it}	-0.414	-3.417^a
	(1.253)	(0.594)
Obs.	4,194	4,194

Notes: a denotes 1% significance, b denotes 5% significance, c denotes 10% significance. Correlations are OLS estimates. Standard errors are in parentheses. Included controls are $\ln L_{\text{aff}}$ and $\ln L_{\text{par}}$.

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

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

we can write

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

[57]The regressions in Table A.1 should not be interpreted as a formal first stage of a Two-Stage Least Squares estimation. The model being 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. These regressions are just computed as a representation of the correlation present in our data between r_{ijt} and r_{i0t}, on one side, and $\text{UCRD}_{it} \times \text{IPR}_{n_{ij}t}$ and UCRD_{it}, on the other side.

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

$$\mathbb{E}[\eta_{ijt}|k_{ijt}, \psi_{ijt-1}, r_{ijt-1}, r_{i0t-1}, \mu_{n_{ij}t}, \kappa_{n_{ij}t}, j \in \mathcal{J}_{it-1}, j \in \mathcal{J}_{it}] = 0. \tag{A.4}$$

Section 4.3 shows that, whether or not we assume that entry decisions are instantaneous, (A.4) implies

$$\mathbb{E}[\eta_{ijt}|k_{ijt}, \psi_{ijt-1}, r_{ijt-1}, r_{i0t-1}, \mu_{n_{ij}t}, \kappa_{n_{ij}t}, j \in \mathcal{J}_{it}] = 0. \tag{A.5}$$

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.4) becomes a function of $\mathbf{S}_{it}$ and, from (10), implicitly a function of k_{ijt}, $\kappa_{n_{ij}t}$, 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}, $\mu_{n_{ij}t}$ 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.4).

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 $\overline{\psi}_{ijt}$ such that, if $\psi_{ijt} \leq \overline{\psi}_{ijt}$, affiliate j is not integrated in multinational i at period t, and the opposite is true if $\psi_{ijt} \geq \overline{\psi}_{ijt}$. The threshold value $\overline{\psi}_{ijt}$ 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.5) as

$$\mathbb{E}[\eta_{ijt}|k_{ijt}, \psi_{ijt-1}, r_{ijt-1}, r_{i0t-1}, \mu_{n_{ij}t}, \kappa_{n_{ij}t}, \psi_{ijt} \geq \bar{\psi}_{ijt}] = 0. \tag{A.6}$$

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}, $\gamma_{n_{ij}t}$. 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.6) as

$$\mathbb{E}[\eta_{ijt}|\psi_{ijt} \geq \bar{\psi}_{ijt}] = 0. \tag{A.7}$$

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

$$\mathbb{E}[\eta_{ijt}|\rho\psi_{ijt-1} + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \mu_{n_{ij}t} + \eta_{ijt} \geq \bar{\psi}_{ijt}] = 0. \tag{A.8}$$

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 $\boldsymbol{\beta}$ 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

$$\iota = \iota_1 \iota_2, \qquad \text{with} \qquad \iota_1 = \frac{1}{\alpha_m} \qquad \text{and} \qquad \iota_2 = \frac{\alpha_m \frac{\sigma-1}{\sigma}}{1 - \alpha_m \frac{\sigma-1}{\sigma}}. \tag{A.9}$$

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 $\sigma_{n_{ij}t}$, and similarly denote the market-year varying parameter ι_2 as $\iota_{2n_{ij}t}$.

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 $\iota_{2n_{ij}t}$ 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 $\{\hat{\iota}_{2n_{ij}t}\}$, and a vector of parameters that do not vary by market or year; (3) estimating this vector of parameters.

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

$$\alpha_m \frac{\sigma_{n_{ij}t} - 1}{\sigma_{n_{ij}t}} = \frac{P^m_{n_{ij}t} M_{ijt}}{Y^*_{ijt}} = \frac{Y^*_{ijt} - VA^*_{ijt}}{Y^*_{ijt}},$$

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 $Y_{ijt} \equiv Y^*_{ijt} \exp(\varepsilon^y_{ijt})$, we write

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

Taking logs, this becomes

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

or, equivalently,

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

Given the expression for ι_2 in equation (A.9), this may be expressed as

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

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,

$$\mathbb{E}[\varepsilon_{ijt} D_{n_{ij}t}] = \mathbb{E}[\varepsilon^y_{ijt} D_{n_{ij}t}] = 0,$$

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

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

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

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

$$va_{ijt} = \kappa_{n_{ij}t} + \tilde{h}(l_{ijt}, k_{ijt}, \hat{\iota}_{2n_{ij}t}; \tilde{\boldsymbol{\beta}}) + \tilde{\psi}_{ijt} + \varepsilon_{ijt} \tag{A.10}$$

where

$$\tilde{h}(l_{ijt}, k_{ijt}, \hat{\iota}_{2n_{ij}t}; \tilde{\boldsymbol{\beta}}) = \tilde{\beta}_l \hat{\iota}_{2n_{ij}t} l_{ijt} + \tilde{\beta}_{ll} \hat{\iota}_{2n_{ij}t} l_{ijt}^2 + \tilde{\beta}_k \hat{\iota}_{2n_{ij}t} k_{ijt} + \tilde{\beta}_{kk} \hat{\iota}_{2n_{ij}t} k_{ijt}^2 + \tilde{\beta}_{lk} \hat{\iota}_{2n_{ij}t} l_{ijt} k_{ijt}, \tag{A.11}$$

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

$$\tilde{\psi}_{ijt} = \mathbb{E}_{t-1}[\tilde{\psi}_{ijt}] + \tilde{\eta}_{ijt}, \tag{A.12}$$

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

$$\mathbb{E}_{t-1}[\tilde{\psi}_{ijt}] = \tilde{\rho} \hat{\iota}_{2n_{ij}t} \psi_{ijt-1} + \tilde{\mu}_a \hat{\iota}_{2n_{ij}t} r_{ijt-1} + \tilde{\mu}_p \hat{\iota}_{2n_{ij}t} r_{i0t-1} + \tilde{\mu}_{ap} \hat{\iota}_{2n_{ij}t} r_{ijt-1} r_{i0t-1} + \tilde{\mu}_{n_{ij}t}. \tag{A.13}$$

Step 3: Estimating $(\tilde{\boldsymbol{\beta}}, \tilde{\rho}, \tilde{\mu}_a, \tilde{\mu}_p, \tilde{\mu}_{ap}, \{\tilde{\gamma}_{n_{ij}t}\})$. Combining equations (A.10), (A.12), and (A.13), we obtain

$$\begin{aligned} va_{ijt} &= \tilde{h}(l_{ijt}, k_{ijt}, \hat{\iota}_{2n_{ij}t}; \tilde{\boldsymbol{\beta}}) + \tilde{\rho}(va_{ijt-1} - \tilde{h}(l_{ijt-1}, k_{ijt-1}, \hat{\iota}_{2n_{ij}t-1}; \tilde{\boldsymbol{\beta}})) \\ &\quad + \tilde{\mu}_a \hat{\iota}_{2n_{ij}t} r_{ijt-1} + \tilde{\mu}_p \hat{\iota}_{2n_{ij}t} r_{i0t-1} + \tilde{\mu}_{ap} \hat{\iota}_{2n_{ij}t} r_{ijt-1} r_{i0t-1} + \tilde{\gamma}_{n_{ij}t} + \tilde{u}_{ijt}, \end{aligned} \tag{A.14}$$

where the error term is $u_{ijt} = \tilde{\eta}_{ijt} + \varepsilon_{ijt} - \tilde{\rho}\varepsilon_{ijt-1}$, and $\tilde{\gamma}_{n_{ij}t} = \tilde{\mu}_{n_{ij}t} + \kappa_{n_{ij}t} - \tilde{\rho}\kappa_{n_{ij}t-1}$. Given (A.14), we can follow analogous steps to those described in section 4.1 in order to obtain consistent estimates of $(\tilde{\boldsymbol{\beta}}, \tilde{\rho}, \tilde{\mu}_a, \tilde{\mu}_p, \tilde{\mu}_{ap}, \{\tilde{\gamma}_{n_{ij}t}\})$.

Notice that we can use the estimates of the parameter vector $(\tilde{\boldsymbol{\beta}}, \tilde{\rho}, \tilde{\mu}_a, \tilde{\mu}_p, \tilde{\mu}_{ap}, \{\tilde{\gamma}_{n_{ij}t}\})$, those of $\{\hat{\iota}_{2n_{ij}t}\}$ and (A.13) to compute the elasticity of $\tilde{\psi}_{ijt}$ with respect to r_{ijt-1} as

$$\hat{\tilde{\mu}}_a \hat{\iota}_{2n_{ij}t} + \hat{\tilde{\mu}}_{ap} \hat{\iota}_{2n_{ij}t} r_{i0t-1},$$

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

$$\hat{\tilde{\mu}}_p \hat{\iota}_{2n_{ij}t} + \hat{\tilde{\mu}}_{ap} \hat{\iota}_{2n_{ij}t} r_{ijt-1}.$$

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: $\{\hat{\iota}_{2n_{ij}t}\}$. 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 $(\boldsymbol{\beta}, \mu_a, \mu_p, \mu_{ap})$ may vary across market and years in a way that is not constrained by the multiplicative terms $\{\hat{\iota}_{2n_{ij}t}\}$. 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

$$y_{ijt} = \kappa_{n_{ij}t} + h(l_{ijt}, k_{ijt}; \boldsymbol{\beta}) + \psi_{ijt} + \varepsilon_{ijt}, \tag{A.15}$$

where $\kappa_{n_{ij}t} = (1/\sigma)q_{n_{ij}t} + p_{n_{ij}t}$, $\boldsymbol{\beta} = \boldsymbol{\alpha}\iota$, $\psi_{ijt} = \iota(\omega_{ijt} + \xi_{ijt})$ and $\iota = (\sigma - 1)/\sigma$.[58] 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):

$$\begin{aligned} y_{ijt} = h(k_{ijt}, l_{ijt}; \boldsymbol{\beta}) + \rho(y_{ijt-1} - h(k_{ijt-1}, l_{ijt-1}; \boldsymbol{\beta})) &+ \mu_a r_{ijt-1} + \mu_p r_{i0t-1} \\ &+ \mu_{ap} r_{ijt-1} r_{i0t-1} + \gamma_{n_{ij}t} + u_{ijt}, \end{aligned} \tag{A.16}$$

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

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:

$$\min_{L_{ijt}} \mathcal{L}(L_{ijt}) = P^l_{ijt} L_{ijt} + P^k_{ijt} K_{ijt} + \lambda_{ijt}(Q_{ijt} - Q_{ijt}(\cdot)),$$
$$\frac{\partial \mathcal{L}_{it}}{\partial L_{ijt}} = P^l_{ijt} - \lambda_{ijt} \frac{\partial Q_{ijt}(\cdot)}{\partial L_{ijt}} = 0,$$

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

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

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

$$\frac{\partial Q_{ijt}(\cdot)}{\partial L_{ijt}} \frac{L_{ijt}}{Q_{ijt}} = \frac{W^l_{ijt}}{Y^*_{ijt}} \zeta_{ijt},$$

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

[58] In 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 $\kappa_{n_{ij}t}$, $\boldsymbol{\beta}$, ι and ψ_{ijt} even though they express slightly different variables.

rewrite this first order condition with respect to labor as

$$\alpha_l + \alpha_{ll} 2 l_{ijt} + \alpha_{lk} k_{ijt} = \frac{W^l_{ijt}}{Y^*_{ijt}} \zeta_{ijt},$$

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

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

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

$$\beta_l + \beta_{ll} 2 l_{ijt} + \beta_{lk} k_{ijt} = \frac{W^l_{ijt}}{Y_{ijt}} \exp(\varepsilon_{ijt}) \tilde{\zeta}_{ijt}. \tag{A.17}$$

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.17) 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.16). We cannot use NLS to estimate the parameters entering (A.16) 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.14). 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.14), 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.[59]

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

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

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

[59] Notice that the coefficients entering (A.16) 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}.

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.17) 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

$$y_{ijt} = \kappa_{n_{ij}t} + \beta_l l_{ijt} + \beta_k k_{ijt} + \psi_{ijt} + \varepsilon_{ijt}, \tag{A.18}$$

where $\kappa_{n_{ij}t} = (1/\sigma)q_{n_{ij}t} + p_{n_{ij}t}$, $\beta_l = \alpha_l \iota$, $\beta_k = \alpha_k \iota$, $\psi_{ijt} = \iota(\omega_{ijt} + \xi_{ijt})$ 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):

$$\begin{aligned} y_{ijt} = \beta_l l_{ijt} + \beta_k k_{ijt} + \rho(y_{ijt-1} - \beta_l l_{ijt-1} - \beta_k k_{ijt-1}) &+ \mu_a r_{ijt-1} + \mu_p r_{i0t-1} \\ &+ \mu_{ap} r_{ijt-1} r_{i0t-1} + \gamma_{n_{ij}t} + u_{ijt}, \end{aligned} \tag{A.19}$$

with $u_{ijt} = \eta_{ijt} + \varepsilon_{ijt} - \rho\varepsilon_{ijt-1}$, and $\gamma_{n_{ij}t} = \mu_{n_{ij}t} + \kappa_{n_{ij}t} - \rho\kappa_{n_{ij}t-1}$. Following the same steps as in Approach 1 above, we obtain the following expression for the first order condition with respect to labor,

$$\beta_l = \frac{W^l_{ijt}}{Y_{ijt}} \exp(\varepsilon_{ijt})\tilde{\zeta}_{ijt}. \tag{A.20}$$

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.19). As discussed above in Approach 1, obtaining a consistent estimate of $(\beta_l, \beta_k, \rho, \mu_a, \mu_p, \mu_{ap})$ using (A.19) 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.[60] 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):

$$Q_{ijt} = (H(K_{ijt}, L_{ijt}; \boldsymbol{\alpha}))^{1-\alpha_m} (M^*_{ijt})^{\alpha_m} \exp(\omega_{ijt}), \tag{A.21}$$

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, $Y_{ijt} = Y^*_{ijt}\exp(\varepsilon^y_{ijt})$, and value added, $Y_{ijt} = Y^*_{ijt}\exp(\varepsilon^{va}_{ijt})$, we can write our measure of the log of materials use by affiliate j of firm i in period t as

$$m_{ijt} = y_{ijt} - va_{ijt} + \varepsilon^y_{ijt} - \varepsilon^{va}_{ijt}. \tag{A.22}$$

In this case, the revenue function becomes

$$y_{ijt} = \kappa_{n_{ij}t} + h(l_{ijt}, k_{ijt}; \boldsymbol{\beta}) + \beta_m m_{ijt} + \psi_{ijt} + \varepsilon^y_{ijt}, \tag{A.23}$$

where $\kappa_{n_{ij}t} = (1/\sigma)q_{n_{ij}t} + p_{n_{ij}t}$, $\boldsymbol{\beta} = \boldsymbol{\alpha}\iota$, $\beta_m = \alpha_m\iota$, $\psi_{ijt} = \iota(\omega_{ijt} + \xi_{ijt})$ and $\iota = (\sigma-1)/\sigma$. Combining (A.22) and (A.23), we can rewrite revenue as

$$\begin{aligned} y_{ijt} &= \kappa_{n_{ij}t} + h(l_{ijt}, k_{ijt}; \boldsymbol{\beta}) + \beta_m(y_{ijt} - va_{ijt}) + \psi_{ijt} + (1+\beta_m)\varepsilon^y_{ijt} - \beta_m\varepsilon^{va}_{ijt} \\ &= \tilde{\kappa}_{n_{ij}t} + h(l_{ijt}, k_{ijt}; \tilde{\boldsymbol{\beta}}) - (\beta_m/(1-\beta_m))va_{ijt} + \tilde{\psi}_{ijt} + ((1+\beta_m)/(1-\beta_m))\varepsilon^y_{ijt} - (\beta_m/(1-\beta_m))\varepsilon^{va}_{ijt} \end{aligned} \tag{A.24}$$

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

$$\begin{aligned} y_{ijt} = {} & h(k_{ijt}, l_{ijt}; \tilde{\boldsymbol{\beta}}) - (\beta_m/(1-\beta_m))va_{ijt} + \rho(y_{ijt-1} - h(k_{ijt-1}, l_{ijt-1}; \tilde{\boldsymbol{\beta}}) - (\beta_m/(1-\beta_m))va_{ijt-1}) \\ & + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \gamma_{n_{ij}t} + u_{ijt}, \end{aligned} \tag{A.25}$$

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

$$\beta_l l_{ijt} + \beta_{ll} 2 l_{ijt} + \beta_{lk} k_{ijt} = \frac{W^l_{ijt}}{Y_{ijt}} \exp(\varepsilon^y_{ijt})\tilde{\zeta}_{ijt}, \tag{A.26}$$

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

[60] Even 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.

materials is

$$\beta_m = \frac{M_{ijt}}{Y_{ijt}} \exp(\varepsilon^y_{ijt})\tilde{\zeta}_{ijt} = \frac{Y_{ijt}\exp(-\varepsilon^y_{ijt}) - VA_{ijt}\exp(-\varepsilon^{va}_{ijt})}{Y_{ijt}\exp(-\varepsilon^y_{ijt})}\tilde{\zeta}_{ijt},$$

or, equivalently,

$$1 - \beta_m \frac{1}{\tilde{\zeta}_{ijt}} = \frac{VA_{ijt}}{Y_{ijt}} \exp(\varepsilon^y_{ijt} - \varepsilon^{va}_{ijt}). \tag{A.27}$$

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.25) as estimating equation. As discussed already above in Approach 1, deriving orthogonality restrictions from (A.25) for estimation requires instruments for y_{ijt-1} and potentially also for l_{ijt}. In addition, identifying β_m using (A.25) 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.[61] 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.26). As this equation is identical to (A.17), 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, $(\beta_l, \beta_{ll}, \beta_{lk})$.

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:

$$Q_{ijt} = (H(K_{ijt}, L_{ijt}, M^*_{ijt}; \boldsymbol{\alpha}))\exp(\omega_{ijt}), \tag{A.28}$$

where

$$H(K_{ijt}, L_{ijt}, M^*_{ijt}; \boldsymbol{\alpha}) \equiv \exp(h(K_{ijt}, L_{ijt}, M^*_{ijt}; \boldsymbol{\alpha})) \tag{A.29}$$

$$\begin{aligned} h(k_{ijt}, l_{ijt}, m^*_{ijt}; \boldsymbol{\alpha}) \equiv{} & \alpha_l l_{ijt} + \alpha_k k_{ijt} + \alpha_{ll} l^2_{ijt} + \alpha_{kk} k^2_{ijt} + \alpha_{lk} l_{ijt} k_{ijt} \\ & + \beta_m m^*_{ijt} + \beta_{mm} m^{*2}_{ijt} + \beta_{lm} l_{ijt} m^*_{ijt} + \beta_{km} k_{ijt} m^*_{ijt}. \end{aligned} \tag{A.30}$$

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, $Y_{ijt} = Y^*_{ijt}\exp(\varepsilon^y_{ijt})$, and value added, $VA_{ijt} = VA^*_{ijt}\exp(\varepsilon^{va}_{ijt})$, we can write the log of measured materials used by affiliate j of firm i in period t as

$$m_{ijt} = y_{ijt} - m_{ijt} = y^*_{ijt} - va^*_{ijt} + \varepsilon^y_{ijt} - \varepsilon^{va}_{ijt} = m^*_{ijt} + \varepsilon^y_{ijt} - \varepsilon^{va}_{ijt}. \tag{A.31}$$

[61] If 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.25). The variable va_{ijt} would still be correlated with u_{ijt} through the period-t firm performance shock η_{ijt}.

In this case, the revenue function becomes

$$y_{ijt} = \kappa_{n_{ij}t} + h(l_{ijt}, k_{ijt}, m^*_{ijt}; \boldsymbol{\beta}) + \psi_{ijt} + \varepsilon^y_{ijt}, \tag{A.32}$$

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

$$\begin{aligned} y_{ijt} &= \kappa_{n_{ij}t} + h(l_{ijt}, k_{ijt}, m_{ijt} - \varepsilon^y_{ijt} + \varepsilon^{va}_{ijt}; \boldsymbol{\beta}) + \psi_{ijt} + \varepsilon^y_{ijt}, \\ &= \kappa_{n_{ij}t} + h(l_{ijt}, k_{ijt}, y_{ijt} - va_{ijt} - \varepsilon^y_{ijt} + \varepsilon^{va}_{ijt}; \boldsymbol{\beta}) + \psi_{ijt} + \varepsilon^y_{ijt}. \end{aligned} \tag{A.33}$$

Combining (8), (9), and (A.33) we obtain

$$\begin{aligned} y_{ijt} &= h(l_{ijt}, k_{ijt}, y_{ijt} - va_{ijt} - \varepsilon^y_{ijt} + \varepsilon^{va}_{ijt}; \boldsymbol{\beta}) \\ &\quad + \rho(y_{ijt-1} - h(l_{ijt-1}, k_{ijt-1}, y_{ijt-1} - va_{ijt-1} - \varepsilon^y_{ijt-1} + \varepsilon^{va}_{ijt-1}; \boldsymbol{\beta})) \\ &\quad + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \gamma_{n_{ij}t} + u_{ijt}, \end{aligned} \tag{A.34}$$

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

Using (A.34) 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} 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.31); therefore, measurement error in sales revenue ε^y_{ijt} 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 $(\varepsilon^y_{ijt}, \varepsilon^y_{ijt-1})$ and value added $(\varepsilon^{va}_{ijt}, \varepsilon^{va}_{ijt-1})$ enters nonlinearly in (A.34). Estimating the parameter vector $(\boldsymbol{\beta}, \rho, \mu_a, \mu_p, \mu_{ap})$ 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.28), (A.29), and (A.30), the profit function in (12), and the assumption 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

$$\beta_l + \beta_{ll} 2 l_{ijt} + \beta_{lk} k_{ijt} + \beta_{lm}(y_{ijt} - m_{ijt} - \varepsilon^y_{ijt} + \varepsilon^{va}_{ijt}) = \frac{W^l_{ijt}}{Y_{ijt}} \exp(\varepsilon^y_{ijt}), \tag{A.35}$$

and

$$\beta_m + \beta_{mm} 2(y_{ijt} - m_{ijt} - \varepsilon^y_{ijt} + \varepsilon^{va}_{ijt}) + \beta_{mk} k_{ijt} + \beta_{lm} l_{ijt} = \frac{Y_{ijt}\exp(-\varepsilon^y_{ijt}) - VA_{ijt}\exp(-\varepsilon^{va}_{ijt})}{Y_{ijt}\exp(-\varepsilon^y_{ijt})}. \tag{A.36}$$

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

$$\ln\left(\beta_l + \beta_{ll} 2 l_{ijt} + \beta_{lk} k_{ijt} + \beta_{lm}(y_{ijt} - m_{ijt} - \varepsilon^y_{ijt} + \varepsilon^{va}_{ijt})\right) = w^l_{ijt} - y_{ijt} + \varepsilon^y_{ijt}, \tag{A.37}$$

and

$$\ln\left(\beta_m + \beta_{mm}2(y_{ijt} - m_{ijt} - \varepsilon^y_{ijt} + \varepsilon^{va}_{ijt}) + \beta_{mk}k_{ijt} + \beta_{lm}l_{ijt}\right)$$
$$= \ln(Y_{ijt}\exp(-\varepsilon^y_{ijt}) - VA_{ijt}\exp(-\varepsilon^{va}_{ijt})) - y_{ijt} + \varepsilon^y_{ijt}. \quad \text{(A.38)}$$

It is not possible to derive from (A.37) and (A.38) moment conditions analogous to those in (15). The reason is that the measurement errors affecting revenue and value added, ε^y_{ijt} and ε^{va}_{ijt}, enter nonlinearly on both the left and right sides of (A.37) and (A.38). 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. $\varepsilon^y_{ijt} = \varepsilon^{va}_{ijt} = 0$ 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.37) and (A.38), 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^y_{ijt} and ε^{va}_{ijt}, 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.34). The terms ε^y_{ijt} and ε^{va}_{ijt} also enter nonlinearly in the function $h(\cdot)$ and, therefore, in (A.34). 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:

$$y_{ijt} = h(l_{ijt}, k_{ijt}, y_{ijt} - va_{ijt}; \boldsymbol{\beta}) + \rho(y_{ijt-1} - h(l_{ijt-1}, k_{ijt-1}, y_{ijt-1} - va_{ijt-1}; \boldsymbol{\beta}))$$
$$+ \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \gamma_{n_{ij}t} + u_{ijt}, \quad \text{(A.39)}$$

where $u_{ijt} = \eta_{ijt}$, and $\gamma_{n_{ij}t} = \mu_{n_{ij}t} + \kappa_{n_{ij}t} - \rho\kappa_{n_{ij}t-1}$. 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

$$\mathbb{E}_{t-1}[\psi_{ijt}] = \rho_1\psi_{ijt-1} + \rho_2\psi^2_{ijt-1} + \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \mu_{n_{ij}t}. \quad \text{(A.40)}$$

Combining (6), (8), and (A.40), we obtain

$$va_{ijt} = h(k_{ijt}, l_{ijt}; \boldsymbol{\beta}) + \rho_1(va_{ijt-1} - h(k_{ijt-1}, l_{ijt-1}; \boldsymbol{\beta})) + \rho_2(va_{ijt-1} - h(k_{ijt-1}, l_{ijt-1}; \boldsymbol{\beta}) - \kappa_{n_{ij}t-1})^2$$
$$+ \mu_a r_{ijt-1} + \mu_p r_{i0t-1} + \mu_{ap} r_{ijt-1} r_{i0t-1} + \gamma_{n_{ij}t} + u_{ijt}, \quad \text{(A.41)}$$

where $u_{ijt} = \eta_{ijt} + \varepsilon_{ijt} - \rho\varepsilon_{ijt-1}$, and $\gamma_{n_{ij}t} = \mu_{n_{ij}t} + \kappa_{n_{ij}t} - \rho\kappa_{n_{ij}t-1}$. As it is clear from (A.41), using this equation to estimate the parameter vector of interest, $(\boldsymbol{\beta}, \rho_1, \rho_2, \mu_a, \mu_p, \mu_{ap}, \{\gamma_{n_{ij}t}\}, \{\kappa_{n_{ij}t}\})$, requires estimating also a large set of market-year fixed effects $\{\kappa_{n_{ij}t-1}\}$ 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, $\{\gamma_{n_{ij}t}\}$. Therefore, from (16), we can write

$$\begin{aligned}\eta'_{ijt} = \widehat{va}'_{ijt} - \beta_k k'_{ijt} - \beta_{kk}(k^2_{ijt})' - \rho(\widehat{va}'_{ijt-1} - \beta_k k'_{ijt-1} - \beta_{kk}(k^2_{ijt-1})') \\ - \mu_a r'_{ijt-1} - \mu_p r'_{i0t-1} - \mu_{ap}(r_{ijt-1}r_{i0t-1})',\end{aligned} \tag{A.42}$$

where $(k^2_{ijt})'$, $(k^2_{ijt-1})'$ and $(r_{ijt-1}r_{i0t-1})'$ denote the residuals from the projection of the variables k^2_{ijt}, k^2_{ijt-1} and $r_{ijt-1}r_{i0t-1}$, respectively, on a full set of market-year fixed effects $\{\gamma_{n_{ij}t}\}$.

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

$$(\hat{\beta}_k, \hat{\beta}_{kk}, \hat{\rho}, \hat{\mu}_a, \hat{\mu}_p, \hat{\mu}_{ap}) = \min_{(\beta_k,\beta_{kk},\rho,\mu_a,\mu_p,\mu_{ap})} \sum_{i,j,t} \{\mathbb{1}\{j \in \mathcal{J}_{it}, j \in \mathcal{J}_{it-1}\} \times \eta'_{ijt}\}, \tag{A.43}$$

where (A.42) 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.42) to define the GMM estimator for the parameter vector $(\beta_k, \beta_{kk}, \rho, \mu_a,$ μ_p, $\mu_{ap})$ described in section 4.2 and in Appendix A.3. Specifically, we compute the optimal two-step GMM estimator based on the following set of unconditional moments

$$\mathbb{E}\left[\eta'_{ijt} \times \begin{pmatrix} k'_{ijt} \\ (k^2_{ijt})' \\ k'_{ijt-1} \\ (k^2_{ijt-1})' \\ \text{UCRD}'_{it-1} \\ (\text{UCRD}_{it-1} \times \text{IPR}_{n_{ij}t-1})' \end{pmatrix} \times \mathbb{1}\{j \in \mathcal{J}_{it}, j \in \mathcal{J}_{it-1}\}\right] = 0,$$

where UCRD'_{it-1} and $(\text{UCRD}_{it-1} \times \text{IPR}_{n_{ij}t-1})'$ denote the residuals from the projection of the variables UCRD_{it-1} and $\text{UCRD}_{it-1} \times \text{IPR}_{n_{ij}t-1}$, respectively, on a full set of market-year fixed effects $\{\gamma_{n_{ij}t}\}$.

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}, r_{ijt} and intrafirm trade im_{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

$$\begin{aligned}\psi_{ij} &\equiv \mathbb{E}\big[\lim_{s\to\infty} \psi_{ijs} \big| r_{i0t}, r_{ijt}, im_{ijt}\big] \\ &= \sum_{s>t} \rho^{s-t}\mu_{n_{ij}s} + \frac{1}{1-\rho} g(r_{ijt}, r_{i0t}, im_{ijt}),\end{aligned} \tag{A.44}$$

where r_{i0t}, r_{ijt}, im_{ijt} are observable and $g(r_{ijt}, r_{i0t}, im_{ijt}) \equiv \mathbb{E}_{t-1}[\psi_{ijt}] - \rho\psi_{ijt-1} - \mu_{n_{ij}t}$ with $\mathbb{E}_{t-1}[\psi_{ijt}]$ defined in (18). In order to derive (A.44) we have applied the property that $\mathbb{E}[\eta_{ijs}|r_{ijt}, r_{i0t}, im_{ijt}] = 0$ 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 both affiliate R&D and imports from parent remain at their period t levels yields

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

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 both parent R&D and affiliate imports from parent remain at their period t levels yields

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

and, finally, setting to zero the contribution of intrafirm imports,

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

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

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.44) above. From (19), the expected long-run productivity the parent firm of multinational i is

$$\psi_{i0} = \mathbb{E}\big[\lim_{s\to\infty} \psi_{i0s} \big| r_{i0t}\big] = \sum_{s>t} \rho^{s-t} \mu_{n_{i0}t} + \frac{1}{1-\rho_0} (\mu_0 r_{i0t} + \mu_{a0} \sum_j r_{ijt-1}). \tag{A.45}$$

Comparing equations (A.44) and (A.45), 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, $\mu_{n_{i0}t}$, as well as those unobserved factors that affect the evolution of performance for each affiliate j, $\mu_{n_{ij}t}$. Being able to identify these parameters would require data on the price index, $P_{n_{ij}t}$, the quantity index, $Q_{n_{ij}t}$, and the price of materials, $P^m_{n_{ij}t}$, in every market and year in which either the parent or an affiliate operates. Such data is not available to us; therefore, Figure 3 reports the distribution of the performance of every affiliate relative to its parent that is exclusively due to the distribution of either R&D spending or intra-firm flows within the multinational firm. Specifically, it provides the value of each percentile of the distribution of

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

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 (21) 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 (20) as

$$
\begin{aligned}
GR_{i0t} &= \mathbb{E}_t\Big[\sum_{s>t}\sum_{j\in\mathcal{J}_{is}}\frac{\partial VA^*_{ijs}}{\partial R_{i0t}}\Big]\\
&= \mathbb{E}_t\Big[\sum_{s>t}\sum_{j\in\mathcal{J}_{is}}\frac{\partial \Psi_{ijt+1}}{\partial R_{i0t}}\frac{\partial VA^*_{ijs}}{\partial \Psi_{ijt+1}}\Big]\\
&= \mathbb{E}_t\Big[\sum_{s>t}\sum_{j\in\mathcal{J}_{is}}\frac{\partial \Psi_{ijt+1}}{\partial R_{i0t}}\frac{\partial \Psi_{ijs}}{\partial \Psi_{ijt+1}}\frac{\partial VA^*_{ijs}}{\partial \Psi_{ijs}}\Big]\\
&= \mathbb{E}_t\Big[\sum_{s>t}\Big[\frac{\partial \Psi_{i0t+1}}{\partial R_{i0t}}\frac{\partial \Psi_{i0s}}{\partial \Psi_{i0t+1}}\frac{\partial VA^*_{i0s}}{\partial \Psi_{i0s}}+\sum_{\substack{j\in\mathcal{J}_{is}\\ j\neq 0}}\frac{\partial \Psi_{ijt+1}}{\partial R_{i0t}}\frac{\partial \Psi_{ijs}}{\partial \Psi_{ijt+1}}\frac{\partial VA^*_{ijs}}{\partial \Psi_{ijs}}\Big]\Big]\\
&= \mathbb{E}_t\Big[\sum_{s>t}\Big[\frac{\partial \psi_{i0t+1}}{\partial r_{i0t}}\rho_0^{s-t-1}\frac{VA^*_{i0s}}{R_{i0t}}+\sum_{\substack{j\in\mathcal{J}_{is}\\ j\neq 0}}\frac{\partial \psi_{ijt+1}}{\partial r_{i0t}}\rho^{s-t-1}\frac{VA^*_{ijs}}{R_{i0t}}\Big]\Big], \qquad \text{(A.46)}
\end{aligned}
$$

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

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

Furthermore, assuming the specifications of the stochastic process of productivity given by (18) and (19), it will be true that

$$
\frac{\partial \psi_{i0t+1}}{\partial r_{i0t}} = \mu_0, \qquad \text{(A.47a)}
$$

$$
\frac{\partial \psi_{ijt+1}}{\partial r_{i0t}} = \mu_p + \mu_{ap}r_{ijt} + \mu_{pm}im_{ijt} + \mu_{apm}r_{ijt}im_{ijt}. \qquad \text{(A.47b)}
$$

Plugging these expressions into (A.46), we obtain

$$
GR_{i0t} =
$$

$$
\mathbb{E}_t\Big[\sum_{s>t}\Big[\mu_0\rho_0^{s-t-1}\frac{VA^*_{i0s}}{R_{i0t}}+\sum_{j=1}^{J_{is}}(\mu_p + \mu_{ap}r_{ijt} + \mu_{pm}im_{ijt} + \mu_{apm}r_{ijt}im_{ijt})\rho^{s-t-1}\frac{VA^*_{ijs}}{R_{i0t}}\Big]\Big]. \qquad \text{(A.48)}
$$

Finally, assuming that $\mathcal{J}_{is}$, VA^*_{i0s} and VA^*_{ijs} remain constant at their period t values for every year $s>t$

$$
GR_{i0t} = \frac{\mu_0}{1-\rho_0}\tau_{i0t}\frac{VA^*_{i0t}}{R_{i0t}} + \sum_{\substack{j\in\mathcal{J}_{it}\\ j\neq 0}}\frac{\mu_p + \mu_{ap}r_{ijt} + \mu_{pm}im_{ijt} + \mu_{apm}r_{ijt}im_{ijt}}{1-\rho}\frac{VA^*_{ijt}}{R_{i0t}}, \qquad \text{(A.49)}
$$

with corresponds to (21) in the main text.

From (A.48), we can also compute the derivative of GR_{i0t} with respect to R_{i0t},

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

and, therefore,

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

or, equivalently,

$$\frac{\partial GR_{i0t}}{\partial R_{i0t}} = \mathbb{E}_t\Big[\sum_{s>t}\Big[\frac{\partial \psi_{i0t+1}}{\partial r_{i0t}}\rho_0^{s-t-1}\frac{VA^*_{i0s}}{(R_{i0t})^2}\Big(\frac{\partial va^*_{i0s}}{\partial r_{i0t}} - 1\Big) + \sum_{\substack{j\in\mathcal{J}_{is}\\ j\neq 0}}\frac{\partial \psi_{ijt+1}}{\partial r_{i0t}}\rho^{s-t-1}\frac{VA^*_{ijs}}{(R_{i0t})^2}\Big(\frac{\partial va^*_{ijs}}{\partial r_{i0t}} - 1\Big)\Big]\Big].$$

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.47a) and (A.47b), the expression simplifies to

$$\begin{aligned}\frac{\partial GR_{i0t}}{\partial R_{i0t}} &= \mathbb{E}_t\Big[\sum_{s>t}\Big[\mu_0\rho_0^{s-t-1}\frac{VA^*_{i0s}}{(R_{i0t})^2}\Big(\mu_0\rho_0^{s-t-1} - 1\Big)\\ &+ \sum_{\substack{j\in\mathcal{J}_{is}\\ j\neq 0}}(\mu_p + \mu_{ap}r_{ijt} + \mu_{pm}im_{ijt} + \mu_{apm}r_{ijt}im_{ijt})\rho^{s-t-1}\frac{VA^*_{ijs}}{(R_{i0t})^2}\\ &\times\Big((\mu_p + \mu_{ap}r_{ijt} + \mu_{pm}im_{ijt} + \mu_{apm}r_{ijt}im_{ijt})\rho^{s-t-1} - 1\Big)\Big]\Big].\end{aligned}$$

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

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 (18),

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

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-US 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.

Table A.2: 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 $\beta_l + \beta_{ll} 2 l_{ijt} + \beta_{lk} k_{ijt}$; Capital Elasticity is the average value of $\beta_k + \beta_{kk} 2 k_{ijt} + \beta_{lk} l_{ijt}$. 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.