Journal of Economics & Management Strategy

Behind the Scenes: Sources of Complementarity in R&D

Marco Ceccagnoli
Scheller College of Business, Georgia Institute of Technology, Atlanta, Georgia
marco.ceccagnoli@scheller.gatech.edu

Matthew J. Higgins
Scheller College of Business, Georgia Institute of Technology, Atlanta, Georgia; and
National Bureau of Economic Research, Cambridge, Massachusetts
matt.higgins@scheller.gatech.edu

Vincenzo Palermo
Scheller College of Business, Georgia Institute of Technology; and
Munk School of Global Affairs, University of Toronto, Toronto Ontario
vincenzo.palermo@utoronto.ca

Management consultants increasingly recommend that internal R&D be outsourced; however, little is known about the substitution or complementarity between internal and external R&D. Through structural estimation of a flexible innovation production function we provide a deeper understanding of firm-level drivers of complementarity between these two types of investments. Our analysis is based on a unique panel data set on the R&D and in-licensing expenditures of pharmaceutical firms. Our results suggest that internal R&D and in-licensing are neither complements nor substitutes. We find that the degree of complementarity is enhanced for firms with stronger absorptive capacity, economies of scope, and licensing experience.

1. Introduction

Markets for technology have been extensively studied; however, there still remains little evidence about the determinants of technology demand in terms of the relationship between internal and external R&D (Arora and Gambardella, 2010). Firms choose their level of integration within the value chain, but the extent to which they adopt different R&D strategies as substitutes or complements remains uncertain. Some firms, such as Morgan Stanley, have advocated a radical shift for the management of R&D in certain

We thank Rajshree Agarwal, Kazuhiro Asakawa, Dan Breznitz, Annamaria Conti, Ronnie Chatterji, Chris Forman, Alfonso Gambardella, John Haltiwanger, Magnus Holmén, Keld Laursen, Anne Miner, Alex Oettl, Maija Renko, Daniel Spulber, Marie and Jerry Thursby, Arvids Ziedonis, seminar participants at Georgia Tech, CCC doctoral consortium (MIT), Academy of Management annual meeting, DRUID conference, and Trans-Atlantic Doctoral Conference (London Business School) for helpful comments and suggestions. We are indebted to Deloitte ReCap for access to their data. We also thank IMS Health Incorporated for their generous support and access to their data. The statements, findings, conclusions, views, and opinions contained and expressed herein are not necessarily those of IMS Health Incorporated or any of its affiliated or subsidiary entities. The statements, findings, conclusions, views, and opinions contained and expressed in this article are based in part on data obtained under license from the following IMS Health Incorporated or affiliate information service(s): IMS MIDAS, June 1997 to December 2008, IMSA Health Incorporated or its affiliates. Higgins acknowledges support from The Imlay Professorship. Ceccagnoli acknowledges funding from the Kauffman Foundation—Roadmap for an Entrepreneurial Economy. Palermo acknowledges funding from the Kauffman Dissertation Program Fellowship. Authors are listed alphabetically and the usual disclaimers apply.

© 2014 Wiley Periodicals, Inc.

industries (Morgan Stanley, 2010; Tollman, 2011). In particular, they argue that the pharmaceutical industry should abandon its current R&D model and fully adopt a "search and development" (S&D) model. Under an S&D framework, firms would abandon all internal research and focus solely on development. Thus, 100% of a firm's drug candidates would come from external licensing. Although full adoption of an S&D model is an extreme position, some pharmaceutical companies have openly acknowledged a move toward more frequent engagement in external licensing. For example, in 2009, GlaxoSmithKline terminated its legendary neuroscience program in order to free up capital to meet its stated goal of allocating 50% of its R&D budget to external projects (Knowles and Higgins, 2011).

The S&D model implicitly suggests that internal and external R&D are substitute activities in the sense that implementation of one activity reduces marginal return on the other activity. Complementarity would arise if an increase in one of these activities increased the marginal returns from the other activity (Milgrom and Roberts, 1990). Substitution between these activities is consistent with the extreme case of backward integration, whereby firms rely exclusively on internal R&D investments. Backward integration dominated the organization of R&D in the past century. Substitution is also consistent with the opposite case, whereby a nonintegrated firm relies exclusively on external technology, perhaps yet to be developed, as in the case of the S&D model. Ultimately, the decision to choose between these two types of R&D is influenced by whether synergies exist between them. For example, internal R&D and licensing could fulfill quite distinct yet complementary purposes. R&D can serve functions not directly tied to the creation of new products, such as concept exploration, hypothesis testing, and market credibility, which are all activities that can complement the investment made on a technology licensed from other firms or institutions.

Our review of the literature suggests that empirical evidence does not conclusively support substitution or complementarity across all industry settings. Moreover, there is surprisingly little research on the contextual factors that determine whether these two activities are complements or substitutes. Accordingly, the major objective of this paper is to provide a deeper understanding of the firm-level drivers that determine the degree of complementarity between internal and external R&D. To accomplish our goal, we provide structural estimates of the degree of complementarity or substitutability between these two types of R&D investments that vary across firms.

Our study is focused on the global pharmaceutical industry, which is an ideal research setting for several reasons. In the pharmaceutical industry, internal productivity failures and the lack of capabilities in emerging technology, coupled with an increase in new external opportunities, have influenced the balance between internal R&D and in-licensing strategies (Malerba and Orsenigo, 2000). Furthermore, internal and external R&D are considered major drivers of firm performance (Scherer, 2007). Finally, the detailed availability of longitudinal measures relating to both internal and external research activities and their product innovation output allows us to directly analyze the marginal productivity of these investments and drivers.

Our results suggest that, on average, internal R&D and in-licensing investments are neither complements nor substitutes in the pharmaceutical industry. However, we show that the degree of complementarity is enhanced for firms with stronger absorptive capacity, economies of scope, and past licensing experience. Taken together, our results highlight the complexity of this relation and suggest that a simple categorization (complement or substitute) may be misleading. In such a context, the framework presented in this paper appears to be valuable, because it recognizes the importance of

heterogeneity across firms in terms of affecting complementarity of internal and external R&D capabilities within narrowly defined industries. Conditional on data availability, such an approach could be readily applied by other researchers to examine similar issues within various industry contexts.

The remainder of the paper is organized as follows: Section 2 discusses the relevant literature and theoretical framework, Section 3 introduces our empirical model, Sections 4 and 5 present data and empirical results, respectively, and Section 6 concludes.

2. Literature Review

2.1 Complements or Substitutes?

Firms must continuously invest in the development of new products in order to stay competitive. Sources of innovative knowledge are no longer limited to internal investments, but they include more significant contributions from external sources, such as licensing. The importance of technology licensing has long been recognized in the literature on industrial organization. However, past research on markets for technology has mostly focused on the supply-side drivers of licensing decisions (Bresnahan and Gambardella, 1998; Arora et al., 2001; Arora and Ceccagnoli, 2006). Less attention has been paid to the incentive to buy technology in the market, particularly on the relationship between internal and external R&D (Arora and Gambardella, 2010). This is an important gap in the literature, because technology buyers in most high-tech industries conduct extensive internal R&D, which could alter their external investment strategy. If it does, this creates a potential tension between developing technology internally and obtaining it externally. This tension raises the question of whether internal and external R&D investments are complements or substitutes. Although a few studies have recently attempted to address this question, results to date are not conclusive.

Several empirical studies support the substitution viewpoint. Pisano's (1990) findings suggest that substitution is driven by transaction costs and their influence on the decision to externally expand R&D. Laursen and Salter (2006) find that internal R&D investments negatively moderate the relationship between external knowledge (licensing) and innovation performance. In a study on investments in advanced Internet technologies, Forman et al. (2008) find a substitute relationship between internal firm resources (e.g., programmers) and external technologies. In a model of technology adoption, they find that the marginal contribution of internal firm resources tends to diminish within large urban areas. It is, therefore, possible that the external resources available in cities are partial substitutes for both establishment-level and firm-level internal resources.

Complementarity between internal and external R&D, on the other hand, implies that these two forms of R&D coexist and are interdependent. Unlike the substitute relationship, complementarity implies that firms acquiring external technologies must also continue to engage in internal R&D. Several studies provide evidence in support of complementarity (Lowe and Taylor, 1998; Glaeser et al., 2000; Cassiman and Veugelers, 2006; Tsai and Wang, 2008). Cassiman and Veugelers (2006), for example, provide empirical evidence in support of complementarity between internal R&D and external technology acquisition strategies (these include licensing, alliances, and acquisitions). In a study of Taiwanese electronics manufacturing, Tsai and Wang (2008) demonstrate that external technology acquisition does not contribute to firm performance *per se*. They find, however, that external acquisition of technology has a positive effect on performance when interacted with internal R&D.

Other empirical evidence, consistent with complementarity, suggests that external know-how can quickly bring new resources to a firm during different stages of production. New knowledge, such as externally generated patents or partially developed compounds, can boost the development process and potentially increase expected revenues. Along these lines, Higgins and Rodriguez (2006) find that internal knowledge is combined with technology acquisition to fill research pipeline gaps. Danzon et al. (2007) argue that firms acquire technology in order to replenish pipeline gaps and respond to excess capacity generated by patent expirations. Similarly, Chan et al. (2007) find that firms engage in the external technology market as a result of downstream cospecialized complementary assets.

In contrast, Vega-Jurado et al. (2009) find no evidence of complementarity nor substitution in the Spanish manufacturing sector. These authors analyze the effect of external knowledge sourcing strategies on the development of both product and process innovation for a sample of innovative Spanish firms. Their results suggest that firms rely on both internal R&D and external knowledge sources, but that the two activities do not have synergistic effects.

In sum, previous research demonstrates the importance of effective internal R&D and external technology acquisition strategies for superior economic performance. However, there is mixed evidence and limited understanding concerning the relationship between these two types of activities, especially their conditioning drivers. Moreover, the scope of prior work has often been limited by data availability, as cross-sectional survey data allow—at best—only analysis of discrete choices of technology that a firm could "make" or "buy" at a specific point in time.

2.2 Drivers of Complementarity

Our reading of the literature suggests that firm-level drivers of the degree of complementarity or substitutability between internal R&D and in-licensing can be grouped into factors determining a firm's absorptive capacity, economies of scope, and licensing experience.

2.2.1 Absorptive Capacity

Absorptive capacity reflects a firm's ability to identify, assimilate, and exploit knowledge from the environment (Cohen and Levinthal, 1989). Arora and Gambardella (1994) formally link this concept to a firm's external technology acquisition strategy. They emphasize two components of absorptive capacity that are relevant to the acquisition of external technology through alliances. One component is the ability to evaluate external technology, which depends on a firm's upstream research capability. Another component is a firm's ability to utilize external technologies, which depends on its technological and development skills.

We build on Arora and Gambardella's contribution by suggesting that both types of firm capabilities tend to be associated with a stronger complementarity between internal and external R&D activities. On one hand, an increase in the cumulated investment in internal R&D, especially when the type of R&D is more basic in nature, tends to generate scientific capabilities, which in turn makes in-licensing more efficient, as it enhances the selection of external technology projects. On the other hand, higher levels of internal R&D, especially when R&D is more geared toward design or development of new products increases the returns from external technology investments by facilitating the effective integration of external technology within the buyer's value chain. We will

exploit this distinction between the ability to evaluate and utilize external technology in our empirical setting in order to guide our empirical measurement and analysis.

2.2.2 Economies of Scope

A second set of drivers of complementarity between internal R&D and in-licensing relate to the concept of economies of scope, defined as the cost savings that are generated from adopting different activities in multiple markets (Panzar and Willig, 1981; Henderson and Cockburn, 1996). The advantage gained through exploitation of economies of scope arises from sharing or jointly utilizing production inputs such as technological resources. When technologies are licensed for use in one market, they can freely or at reduced additional cost be readopted to other markets or products. Therefore, the opportunity to share technologies across different projects facilitate the generation of synergies among them by creating links between resources that would otherwise remain separate.

Although the logic of economies of scope typically refers to the benefits of related diversification in terms of cost advantages, these benefits can also be formulated in terms of products and services. The external knowledge developed for a given technological area may potentially be beneficial to the development of products in other technological areas. Given that knowledge can be articulated and codified within the firm (Zollo and Winter, 2002), the external knowledge acquired for a specific project can be utilized to improve the current development of products in other technological areas.

Following this logic, we expect that firms with broader experience across different technological areas to be characterized by a stronger degree of complementarity between internal R&D and in-licensing. This implies that such firms may be using knowledge developed in different fields additively in the innovative process (Henderson and Cockburn, 1996). Complementarities may arise if technologies purchased from external sources have different technical specifications, and thus are useful to fulfill internal capability gaps. In such cases, economies of scope should increase the synergetic combination of internal and external inputs.

2.2.3 Licensing Experience

The logic underlying the effect of prior licensing experience on the complementarity between internal and external R&D is similar, in many respects, to the concept of absorptive capacity examined earlier. Licensing experience refers to the cumulative experience in leveraging external knowledge, whereas absorptive capacity is based on the cumulative experience developed by investing in internal knowledge. Under this view, collaborative agreements, such as licensing, joint ventures, and acquisitions, may enhance a firm's ability to more effectively combine internal and external technologies.

The literature suggests that firms with prior licensing experience are more likely to have developed effective communication mechanisms, more flexible organizational structures, and other successful organizational routines that can facilitate the integration of external technologies within existing R&D structures (Zollo and Winter, 2002). Indeed, firms vary in the extent to which their organizational structure supports the management of technology acquisition. For example, Pfizer has recently invested in creating a new division called "the Research Network Initiative," tasked to make external technologies from their various partnerships more accessible to internal scientists.[1]

Furthermore, similar to the effect of scientific capabilities highlighted in the previous section, firms with more extensive licensing experience are better able to identify

1. http://www.labnews.co.uk/comment/big-ask/dating-agency-scientists-andrew-mcelroy/. Accessed June 6, 2012.

valuable external technologies that best fit their internal R&D efforts, thus increasing the synergies between the two activities.

3. Model Description and Estimation Procedure

3.1 Translog Innovation Production Function

We assume that at time t each firm j is characterized by an innovation production function (n), which depends on investments for the acquisition of external technology (R_e), internal R&D expenditure (R_i), and a constant term that represents firm-specific effects as well as other exogenous components affecting the productivity of resources invested in innovation (S):

$$n_{jt} = f\left(R_{ijt}, R_{ejt}, S_{jt}\right). \tag{1}$$

Hereafter, the firm and time subscripts are omitted for simplicity.

We adopt a Translog specification, which is the most widely used flexible functional form (Greene, 2012).[2] The Translog specification is defined as follows:

$$n = R_i^{\alpha_i} R_e^{\alpha_e} e^{S+\beta_i(\mathrm{Ln}R_i)^2+\beta_e(\mathrm{Ln}R_e)^2+\gamma_{ie}\mathrm{Ln}R_i\mathrm{Ln}R_e+u}, \tag{2}$$

where R_i and R_e represent internal and external research, respectively; S is the exogenous component of the production function; α_{i}, α_{e}, β_i, β_e, and γ_{ie} are parameters to be estimated, and u is the econometric disturbance. We estimate the Translog by taking logarithms of both sides of equation (2), which allow us to employ linear estimation techniques. Equation (2) shows that it is possible to innovate even if a firm does not invest in these two types of R&D, due to the effect of an exogenous component, S, which might include factors such as knowledge flows from other sources (e.g., firms or universities).

The marginal productivity of internal and external R&D using the Translog specification can be written as follows:

$$\frac{dn}{dR_i} = \frac{n}{R_i}\left(\alpha_i + 2\beta_i \mathrm{Ln}R_i + \gamma_{ie}\mathrm{Ln}R_e\right) = \frac{n}{R_i}Z_i, \tag{3}$$

$$\frac{dn}{dR_e} = \frac{n}{R_e}\left(\alpha_e + 2\beta_e \mathrm{Ln}R_e + \gamma_{ie}\mathrm{Ln}R_i\right) = \frac{n}{R_e}Z_e, \tag{4}$$

where Z_i and Z_e in (3) and (4) represent the elasticity of the number of innovations with respect to internal and external R&D, respectively. We then compute the degree of complementarity or substitutability by estimating the following cross-partial derivative:

$$\frac{d^2n}{dR_i dR_e} = \left[\frac{n}{R_e}\left(\alpha_e + 2\beta_e \mathrm{Ln}R_e + \gamma_{ie}\mathrm{Ln}R_i\right)\frac{Z_i}{R_i} + \frac{n\gamma_{ie}}{R_i R_e}\right] = \frac{n}{R_i R_e}\left(Z_i Z_e + \gamma_{ie}\right) = \frac{n}{R_i R_e}\tilde{Z}, \tag{5}$$

2. We empirically tested the adoption of the Translog functional form by estimating another flexible functional form, the constant elasticity of substitution (CES)-Translog specification (Pollak et al., 1984). This is a functional form that is compatible with a wider range of substitution possibilities than the Translog and which nests Cobb-Douglas, CES, and Translog production functions. By testing its estimated coefficients, we found that both the Cobb-Douglas and the CES specifications were rejected, whereas the Translog was not. Empirical estimations and the coefficient tests are available in the online Appendix. We opted to use the Translog rather than CES-Translog due to the nonlinearities involved in its estimation, which did not allow us to achieve convergence in several of the specifications that we present in this paper.

where $\tilde{Z} = (Z_i Z_e + \gamma_{ie})$ and all other variables are as defined above. In contrast to the signs for CES and Cobb-Douglas, the sign of $\frac{d^2 n}{d R_i d R_e}$ for the Translog functional form is less intuitive. Although the values of R_i, R_e, and n are positive, the sign of $\tilde{Z}$ is ambiguous and we cannot predict *ex ante* whether internal and external R&D investments are complements or substitutes. However, we can estimate the predicted value of n and $\tilde{Z}$ for each firm-year by estimating the log-log specification of the innovation production function expressed by equation (2). We then evaluate $\frac{d^2 n}{d R_i d R_e}$ for representative (mean or median) values of R_i, R_e, and the exogenous predictors of n, with particular attention to the firm-level drivers of complementarity summarized in Section 2.2.

There are several advantages of using a Translog specification to analyze input substitutability (Berndt and Christensen, 1973). First, it is a flexible functional form that is considered a second-order Taylor series approximation to a general, but unknown "true" production function. Second, its flexible functional form allows us to test for complementarity versus substitution using the second-order derivative of innovative output with respect to internal and external R&D. With the Cobb-Douglas, specification inputs are necessarily complements, in the sense that the marginal productivity of any one input increases as the other inputs increase. Moreover, with the Translog functional form we can analyze firm-level drivers of complementarity because input substitutability is allowed to vary with the quantity of inputs as well as other firm characteristics. Third, it represents a specification that is linear in the parameters, which allows estimation through least square methods. Finally, the Translog specification imposes fewer restrictions on the production function and substitution elasticities relative to Cobb-Douglas, thus reducing potential biases in estimated coefficients.

Estimates using the Translog production function should be used with caution, however. Indeed, a common disadvantage of the Translog is the interpretation of the coefficients given the nonlinearity of the relationship between inputs. Because of this, our main focus in the interpretation of the results is the sign of the cross-partial derivatives described later. In addition, the Translog production function with large number of inputs may suffer from collinearity because of the large number of parameters to be estimated.[3] In our model, we limit the number of innovation inputs to two: internal R&D and licensing investments, thus reducing the level of complexity of the model and the potential collinearity between variables. Finally, the Translog production function may not be a satisfactory approximation to the true production function over the full range of input levels considered. The error of approximation to the unknown production function is assumed to be embodied in the econometric disturbance.

3.2 Empirical Strategy

Our estimation procedure involves two steps. First, we estimate the Translog specification modeled by equation (2). Second, we can compute the sign and magnitude of the cross-partial derivative defined in equation (5). An evaluation of this cross-partial derivative suggests that the cross-partial (5) does not necessarily have the same sign of γ_{ie}, for example, the coefficient associated with the interaction between the logs of internal and external R&D.

The production function presented in the previous sections can be used in the context of a profit maximization model with endogenous internal and external R&D

3. If the number of production inputs equals n, the number of estimated parameters is equal to $\frac{n(n+3)}{2}$.

investment levels (available in the online Appendix).[4] Such a model generates exclusion restrictions that imply that variables affecting the optimal level of internal and external R&D do not affect the innovation production function other than through R_i and R_e. This provides information about instrumental variables that can be utilized in order to deal with the endogeneity of internal and external R&D. The source of endogeneity comes from unobserved factors that may drive both the production of innovations as well as the efficiency of internal and external R&D investments. As discussed more fully below, we use exogenous drivers of the expected value of an innovation and organizational factors driving the acquisition of external technology as instruments for internal and external R&D investments in the innovation production function. We also experiment using controls for unobserved firm-specific heterogeneity to test the sensitivity of the results to our identification strategy.

4. Data

Our sample is based on a unique longitudinal data set built from a variety of sources. We began by creating a comprehensive list of global pharmaceutical firms from Pharmaprojects that were active in drug development at any point during 1997–2005. Data include both the timeline of drug development (e.g., the various stages of clinical trials, FDA approval, and project discontinuations) and detailed information on the potential size of the market.

Next, we matched our list of firms with Compustat, collecting data on firm sales, total R&D expenditures, and the number of firm employees. Licensing information was obtained from Deloitte ReCap and includes data on royalties, upfront payments, and milestones. Finally, from IMS MIDAS^{TM} we obtained product-level promotion expenditures as well as number of competitors in each therapeutic category. All financial variables are in year 2000 constant US dollars. Descriptive statistics are provided in Table I and correlations are presented in Table II.

Our final sample consists of 94 pharmaceutical firms active in drug development between 1997 and 2005. Of those, 85% of the firms were located in North America and 12% were located in Europe and the United Kingdom. The average firm has approximately 11 compounds in its pipeline. Our firms, like most major pharmaceutical companies, operate in a number of therapeutic areas. In the sample, the average number of therapeutic categories per firm is six. Almost one-third of the compounds under development are focused in three therapeutic areas: central nervous system, alimentary tract and metabolism, and cardiovascular.

4.1 Dependent Variables

4.1.1 Product Pipeline

Our dependent variable is the firm-year product pipeline, which represents a firm's innovative output. The importance of studying a firm's pipeline is based on the idea that compounds are developed in stages, all of which require different resources and capabilities in order to reach commercialization. These resources can be developed internally or acquired through the markets for technology. Using data from Pharmaprojects, we generate a yearly pipeline stock by cumulating the number of FDA-approved drugs and

4. The model has a structure and assumptions similar to the optimization model of Arora et al. (2008).

Table I.
Descriptive Statistics

Variable	Mean	SD	Min	Max
Product pipeline	1.764	1.184	0	5.029
In-licensing (deflated, Mil. $; stock)	239.443	633.850	0	5,184.333
Internal R&D (deflated, Mil. $; stock)	1,345.112	3,195.205	0.473	28,756.440
Industry promotion (deflated, thousands $)	795,625	747,352.8	0	3,271,951
Trademarks (stock)	11.590	34.003	0	426
Expected market size (deflated, thousands $)	2,158.068	1,556.527	0	10,217.840
Competitors	1,308.882	578.575	201	2,626
Firm size (hundreds)	13.841	26.717	0.001	122
Scientific publications (stock)	450.178	1,232.055	0	11,440
North America	0.849	0.359	0	1
Europe	0.116	0.320	0	1
Other	0.035	0.184	0	1
Number of ATCs	6.781	5.782	1	16
% Licensed compound (Phase 1)	0.029	0.076	0	1
% Licensed compound (Phase 2)	0.055	0.129	0	1
% Licensed compound (Phase 3)	0.056	0.139	0	1
Main therapeutic areas				
ATC A	0.112	0.316	0	1
ATC B	0.023	0.151	0	1
ATC C	0.095	0.294	0	1
ATC D	0.066	0.249	0	1
ATC G	0.050	0.217	0	1
ATC H	0.005	0.072	0	1
ATC J	0.102	0.302	0	1
ATC K	0.009	0.095	0	1
ATC L	0.043	0.203	0	1
ATC M	0.031	0.174	0	1
ATC N	0.145	0.352	0	1
ATC P	0.001	0.036	0	1
ATC R	0.061	0.240	0	1
ATC S	0.030	0.171	0	1
ATC T	0.009	0.095	0	1
ATC V	0.013	0.114	0	1

$N = 767$.

those being developed for each firm in our sample.[5] To account for development uncertainty, compounds are weighted by average probabilities of successfully reaching Food and Drug Administration (FDA) approval, conditional on their phase of development (Grabowski, 2002). In this way, we provide greater weight to later-stage drug candidates (Higgins and Rodriguez, 2006). This is consistent with our objective to compare the efficiency of internal R&D and in-licensing in obtaining new, marketable products.

4.2 Independent Variables

4.2.1 Internal R&D Investments

We compute internal R&D investments using data from Compustat and Deloitte ReCap. R&D data from Compustat include expenditures in R&D that could be performed

5. To deal with observations equal to zero (10% of our sample), we compute our pipeline variable as $\log(1 + x)$. We also tried Poisson's estimation for count data models. Our results remain robust.

Table II.
Correlation Table

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Product pipeline	1												
2. In-licensing	0.496	1											
3. Internal R&D	0.729	0.677	1										
4. Industry promotion	−0.206	−0.017	−0.127	1									
5. Trademarks	0.410	0.548	0.618	−0.019	1								
6. Expected market size	0.365	0.103	0.238	0.019	0.169	1							
7. Competitors	0.086	0.136	0.197	0.491	0.134	0.085	1						
8. Firm size	0.643	0.421	0.756	−0.159	0.509	0.194	0.157	1					
9. Scientific publications	0.501	0.437	0.565	0.058	0.528	0.088	0.188	0.638	1				
10. Number of ATCs	0.660	0.385	0.589	−0.125	0.467	0.219	0.213	0.639	0.473	1			
11. % Licensed compound (Phase 1)	0.163	0.257	0.211	0.016	0.176	0.149	0.129	0.120	0.111	0.100	1		
12. % Licensed compound (Phase 2)	0.042	0.131	0.088	0.072	0.001	−0.166	0.123	0.091	0.081	0.075	0.011	1	
13. % Licensed compound (Phase 3)	0.036	0.091	0.112	0.076	0.050	−0.077	0.076	0.053	0.027	0.053	0.042	0.031	1

internally or externally.[6] In order to isolate internal R&D, we use licensing data from Deloitte Recap and subtract it from the Compustat data. The resulting difference is our proxy for purely internal R&D expenditures. Finally, since developed knowledge can become obsolete over time, we use a 15% depreciation rate to compute an internal R&D stock variable (Hall, 1993).

4.2.3 In-licensing Investments (external R&D)

We use Deloitte ReCap data to collect licensing payments. Our in-licensing variable is based on the sum of milestones and upfront payments.[7] As with the internal R&D variable, we build the stock of licensing investment using a 15% depreciation rate (Hall, 1993). In the case of missing values, we imputed the payments based on the average investment for agreements with similar characteristics, such as the same year of signing, stage at signing, disease, and type of technology.[8] Because the stock of licensing expenditures also captures a firm's licensing experience, we also use this variable to evaluate the extent to which such experience may affect the degree of complementarity between the internal generation and external acquisition of technologies.

Note that due to the log-log transformation necessary to estimate the Translog production function (2), we use the $\log(1 + x)$ transformation for inputs and output with zero values, which may affect our estimates. We will therefore focus our conclusions on the qualitative and statistical significance of our results rather than their magnitude.[9]

4.3 Instrumental Variables for Internal and External R&D

As noted earlier, internal R&D (R_i) and in-licensing (R_e) are correlated with unobserved productivity factors affecting both inputs and output of innovation. Our R&D optimization model (not reported) suggests that variables affecting the expected value of an innovation should only affect the production of innovation through these variables' effects on the levels of internal and external R&D investments.[10] We, therefore, use variables that should affect the profitability of marketed drugs, such as potential size of the

6. In-licensing upfront fees and milestones are expensed when incurred as R&D expenditures (FAS 2R.12). The following examples from public filings explain the underlying accounting principles: (1) ABBOTT 2010 10-K SEC filing (p. 51) states: "Internal research and development costs are expensed as incurred. Clinical trial costs incurred by third parties are expensed as the contracted work is performed. Where contingent milestone payments are due to third parties under research and development arrangements, the milestone payment obligations are expensed when the milestone results are achieved." (2) BIOMARIN 2010 10-K SEC filings (p. 43) states: "Research and development expenses include expenses associated with contract research and development provided by third parties . . . Amounts due under such arrangements may be either fixed fee or fee for service, and may include upfront payments, monthly payments and payments upon the completion of milestones or receipt of deliverables." (3) MERCK 2010 10-K SEC filings (p. 115) states: "Research and development is expensed as incurred. Upfront and milestone payments due to third parties in connection with research and development collaborations prior to regulatory approval are expensed as incurred."

7. We are not able to include royalties in our in-licensing measure because they are included in the income statement as part of operating expenses or as cost of sales and are not explicitly available in a consistent way in either public documents or Deloitte ReCap. Given this limitation, we acknowledge that our in-licensing measure is downward biased and most likely provides a lower bound of the in-licensing effect on innovative output.

8. Only 9% of data had missing values for this variable. To check the robustness of our results to the imputation method, we re-estimated the model without the imputed values, and the results were unchanged.

9. We test the sensitivity of the results to the $\log(1 + x)$ substitutions in several ways. As noted earlier, we estimated a Poisson count data model with fixed effects to deal with firms with zero pipeline products in a given year. We also experimented adding two dummy variables equal to one for observations with no internal or external R&D stocks, as controls. All these robustness analyses indicate that our main results are qualitatively robust.

10. The formal optimization model is available in the online Appendix.

market, number of competitors, industry resources devoted to promotion, and strength of a firm's complementary assets. As shown in the empirical results section, the above instruments appear to have sufficient explanatory power and exploit different sources of variation, as indicated by the tests for instrument validity reported in the tables.

First, we use potential product market size as an instrument for internal and external R&D investments, because it reflects exogenous drivers of the future demand of the firm. In the case of successful approval and commercialization, each firm is able to service the potential market and gain the associated revenues. The larger the expected size of the market, the higher will be the overall R&D effort (both internal and external investments) to develop a final product (Acemoglu and Linn, 2004). Pharmaprojects includes estimates of the potential product market size for drugs in development. We compute the expected market size for pipeline products by summing the estimated values of each firm's drugs in each year.

Second, we use the number of competitors to proxy for the incentive to be innovative and productive. Although the effect on incentives for product innovation is *ex ante* ambiguous, the number of competitors does affect market prices and demand elasticity (Vives, 2008). This variable reflects the number of firms with at least one product sold in the main therapeutic area (ATC) of the focal company.[11] The data were collected from IMS MIDAS™.

Third, we employ detailing expenditures at the industry level, obtained from IMS MIDAS™, to capture the impact of industry resources devoted to promotion on the incentive to innovate. Our *Industry Promotion* variable includes detailing investments defined as promotion activities directed toward physicians and hospitals, journal advertising, and direct-mail and is computed based on the primary ATC of the focal firm.

Finally, we also use the stock of a firm's trademarks to proxy for that firm's brand capital (Fosfuri et al., 2008) and strength of downstream complementary assets, important drivers of the appropriability of returns from innovation (Teece, 1986). This instrument should be correlated with both internal and external R&D investments, because it enhances the appropriability of both types of investments. We collect data on active trademarks from the Unites States Patent and Trademark Office (USPTO) and use them to build a stock variable.

Our identification strategy is based on the assumption that the employed instruments are uncorrelated with the unobserved factors affecting the productivity of a firm's innovative activity. In other words, the chosen instrumental variables should affect the decision to invest in internal and external R&D, but should not independently affect the innovative output of the firm. This can be defended using arguments from economic theory. First, as argued above, economic models based on first principles suggest factors affecting the expected value of an innovation should affect resources devoted to internal and external R&D but should not independently affect the productivity of those activities. Second, our therapeutic category-level instruments are not a choice variable for the firm.

Our instruments based on the trademarks owned by our focal firms, however, may raise endogeneity concerns. Therefore, to further support our identification strategy,

11. ATC stands for Anatomical Therapeutic Chemical as defined by the World Health Organization (http://www.whocc.no/). These therapeutic classes are: A: alimentary tract and metabolism; B, blood and blood forming organs; C, cardiovascular system; D, dermatologicals; G, genitourinary system and sex hormones; H, systemic hormonal preparations, excl. sex hormones and insulins; J, anti-infectives for systemic use; L, antineoplastic and immunomodulating agents; M, musculoskeletal system; N, nervous system; P, antiparasitic products, insecticides, and repellents; and R, respiratory system.

we estimated the *C*-statistic (also known as "generalized method of moments (GMM) distance" or "difference-in-Sargan" statistic) (Hayashi, 2000). The *C*-statistic allows for a test of validity of a subset of instruments, and it is defined as the difference between the Sargan-Hansen statistic of the equation with the set of instruments assumed to be valid (under both the null and the alternative hypothesis) and the same statistic of the equation with all instruments.[12] As shown in our empirical tables, we support the null hypothesis that the set of instruments, which includes *trademarks*, its square term, and its interactions are valid. That said, it is important to remark that identification in our model is achieved through assumptions based on economic reasoning.

4.4 Complementarity Drivers and Other Control Variables

One theoretical argument related to absorptive capacity suggests that the firm's cumulated investment in basic research is complementary to in-licensing. Data pertaining to analysis of this argument are typically unavailable using secondary sources. From an empirical point of view, a way around this problem is to identify the type of R&D conducted by each firm. Indeed, the complementarity between commonly observed measures of R&D (which includes applied R&D activities) should increase the more a firm conducts relatively basic research activities.

Consistent with this idea, Cassiman and Veugelers (2006) suggest that the extent to which a firm relies on more "basic" know-how affects the strength of the complementarity between internal and external innovation strategies. Therefore, in order to measure a focal firm's type of R&D we use their cumulative number of scientific publications, according to data provided by the Web of Science. A strong scientific publications record indicates that a firm's technology is based on advances in science. As a measure of absorptive capacity, which reflects a firm's ability to effectively integrate external technology, we follow Arora and Gambardella (1994) and utilize two alternative measures as a robustness check: the cumulative number of patents granted each year to the focal firm (available from the USPTO) and the cumulative stock of internal R&D.

To measure the potential for economies of scope across different scientific fields, we use the total number of therapeutic areas covered by the drugs in the pipeline of the focal firm each year, which we label, "number of ATCs". Firms that operate in different ATCs may develop capabilities unique to a specific therapeutic area and exploit possible economies of scope. Moreover, innovations in the pipeline can often be used in multiple therapeutic areas, thereby increasing their application possibilities. For example, Topamax® was originally approved as an anti-epileptic but was subsequently used for obesity and peripheral pain.

Among other exogenous variables, we include firm size, which is measured by the total number of firm employees (obtained from Compustat) and intended to control for size-related factors that might drive differences in innovative performance. To control for possible differences in uncertainty between in-licensed and internally developed drugs, we include the percentage of licensed compounds (gathered from Pharmaprojects) that a firm has at each phase of the clinical development process. Indeed, firms that license new compounds may face a higher probability of success because they pay for a compound

12. The *C*-statistic was implemented using the *orthog* option of the Stata command xtivreg2. The null hypothesis is that both the set of valid instruments and the additional set of "suspected instruments" are valid. The *C*-statistic is distributed as chi-squared with the number of instruments tested as degree of freedom.

that has already gone through part of the earliest and more uncertain stages of the development process.[13]

Finally, to control for technological opportunities and other unobserved factors associated with the main technological field of the focal firm, we identify the primary ATC as the therapeutic are with the highest level of annual sales. We include a set of dummy variables that equal one for the main therapeutic area of the focal firm (based on the primary ATC) and zero otherwise. Given our definition of primary ATC, the ATC dummy variables vary over time. We also include specifications with year and firm fixed-effects. The latter are included to control for firm heterogeneity. In models without firm-fixed effects, we also include 4-digit SIC-code and geographic location dummies (e.g., North America, Europe, and other).

5. Results

5.1 Estimating the Degree of Complementarity/Substitutability

We start by estimating the coefficients of the Translog production function (equation (2)) and then estimate the degree of complementarity and its distribution across key firm characteristics. To facilitate estimation and interpretation of the coefficients, we adopt a log-log form of the Translog. This transformation makes the model linear with respect to the natural logarithm of our main independent variables. We then estimate the degree of complementarity by taking the derivative with respect to the logarithm of internal R&D and in-licensing (external R&D) as defined by equation (5).

The estimates of the Translog are reported in Table III. We use three different estimation methods: a benchmark panel data fixed-effects model with instrumental variables estimated with GMM (columns 1–4), a panel random effects model (column 5), and a panel fixed effects model (column 6). The magnitude and significance of the cross-partial derivative $\frac{d^2n}{dR_i dR_e}$ (equation (5)) associated with the models are also reported in Table III. The cross-partials are evaluated at the mean and median values of the variables included in equation (5).

Overall, these results suggest that internal R&D and in-licensing expenditures are neither complements nor substitutes. In particular, the estimated cross-partials presented in Table III are not significantly different from zero across estimation methods.[14]

One possible explanation may reside on the specificity of the drug discovery process. Pharmaceutical firms rely on external technologies in all development stages, and licensed drugs may be used to either substitute an existing stream of research or to complement it. In-licensing is one way to access new knowledge, and new knowledge boosts innovation. However, in-licensing may have two opposite mechanisms. On one hand, external knowledge can fill gaps in internal capabilities. On the other hand, external knowledge can complement internal knowledge by integrating the two sources of knowledge. These results may not be significant because the complementarity effect

13. Pisano (1997) finds evidence of the existence of a market for lemons in the external technology market. If true, this would suggest that firms would not achieve any reductions in risk and the expectations for success of those products would be less than internally developed molecules. However, Arora et al. (2009) find the opposite to be true. They find that compounds licensed during preclinical trials are as likely to succeed as internal compounds of the licensor. Danzon et al. (2005) also find that products developed in an alliance tend to have a higher probability of success.

14. Table AI in the online Appendix replicates the estimates of our benchmark model (column 4, Table III) and compares them to results obtained using different combinations of instrumental variables. Our results are robust to these alternative specifications.

Table III.
Panel Estimates of the Translog Innovation Production Function (full sample)

	Panel GMM FE				Panel RE	Panel FE
Dependent Variable: Ln(1 + Pipeline)	(1)	(2)	(3)	(4)	(5)	(6)
Ln(R&D)	0.298	0.508***	0.516***	0.529**	0.178**	0.209***
	(0.189)	(0.156)	(0.162)	(0.213)	(0.073)	(0.072)
Ln(licensing)	− 0.262	− 0.231*	− 0.318***	− 0.242	0.043	0.007
	(0.162)	(0.121)	(0.104)	(0.153)	(0.045)	(0.042)
(Ln R&D)^2	− 0.0003	− 0.029	− 0.024	0.013	0.003	− 0.009
	(0.034)	(0.025)	(0.026)	(0.037)	(0.008)	(0.008)
(Ln licensing)^2	0.016	0.021	0.024	0.066***	− 0.005	− 0.011*
	(0.026)	(0.021)	(0.016)	(0.024)	(0.005)	(0.006)
Ln(R&D)* Ln(licensing)	− 0.002	0.001	0.0009	− 0.050	− 0.001	0.010
	(0.048)	(0.035)	(0.027)	(0.039)	(0.009)	(0.008)
Scientific publications		− 0.058	− 0.023	− 0.033	0.018	− 0.015
		(0.052)	(0.040)	(0.054)	(0.028)	(0.029)
% Compound licensed—Phase I		0.582*	0.642**	0.684*	0.409*	0.443*
		(0.329)	(0.295)	(0.351)	(0.220)	(0.229)
% Compound licensed—Phase II		0.213	0.376**	0.277	0.147	0.212
		(0.185)	(0.169)	(0.214)	(0.152)	(0.146)
% Compound licensed—Phase III		0.0803	0.102	-0.030	0.287*	0.380***
		(0.144)	(0.194)	(0.259)	(0.148)	(0.144)
Firm size		0.009	0.006	− 0.00002	− 0.0007	− 0.001
		(0.006)	(0.005)	(0.008)	(0.002)	(0.002)
North America					0.453	
					(0.290)	
Europe					1.162***	
					(0.357)	
Number of ATC					0.067***	
					(0.017)	
Firm fixed effect	Yes	Yes	Yes	Yes	No	Yes
ATC dummies	No	No	Yes	Yes	Yes	Yes
Time dummies	No	No	No	Yes	Yes	Yes

Continued

Table III.
Continued

	Panel GMM FE				Panel RE	Panel FE
Dependent Variable: Ln(1 + Pipeline).	(1)	(2)	(3)	(4)	(5)	(6)
Observations	622	622	622	622	767	767
Log-likelihood	−252.6	−192.6	−242.6	−296.3		−106.3
Cluster	72	72	72	72	92	92
Over-identification test (p-value)	0.642	0.189	0.498	0.476		
C-statistic (p-value)	0.972	0.978	0.762	0.272		
Cross-partial						
Mean	−0.005	−0.002	−0.004	−0.01	−0.0001	0.001
	(0.007)	(0.004)	(0.005)	(0.007)	(0.001)	(0.001)
Median	−0.004	−0.001	−0.003	−0.006	−0.0001	0.0008
	(0.005)	(0.003)	(0.004)	(−0.005)	(0.001)	(0.0007)
Endogenous RHS variables	F-statistic (excluded instruments from the Translog production function)					
Ln(RD)	23.64^{***}	5.53^{***}	6.00^{***}	12.28^{***}		
Ln(licensing)	11.99^{***}	3.69^{***}	3.45^{***}	3.01^{**}		
$(\text{Ln R\&D})^2$	31.93^{***}	5.83^{***}	5.67^{***}	12.32^{***}		
$(\text{Ln licensing})^2$	18.98^{***}	8.67^{***}	7.84^{***}	6.78^{***}		
$\text{Ln(R\&D)}^{*}\ \text{Ln(licensing)}$	16.39^{***}	6.59^{***}	5.54^{***}	4.19^{***}		

Notes: Standard errors clustered by firms in parentheses.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The excluded instruments from the main equation (Translog innovation production function) are the following: the natural logs, the square and cross-products of the natural logs of *Industry Promotion*, *Trademarks*, *Potential product market size, and Number of competitors*. ATC dummies vary over time because based on the primary ATC, which is defined in terms of annual sales. The cross-partial derivative presents estimates of $\frac{d^2 n}{d R_i d R_e}$ (equation (5)).

experienced by some firms may be offset by the negative effect experienced by others. Importantly, it follows that studying complementarity without understanding its drivers and the distribution across firms' characteristics may generate misleading results.

The results presented in Table III improve upon the existing literature in several ways. Our use of in-licensing investments provides more direct evidence on the marginal productivity of the financial resources invested in innovation. The extant literature more commonly uses a stock of external deals as a measure of external R&D or self-reported discrete measures of whether a firm acquires technology in the market in cross-sectional settings. Finally, and more substantively, our results indicate that internal R&D and in-licensing do not, *on average*, have a significant *joint* effect on the production of new drugs. However, there are significant differences in the estimated degree of complementarity across firms. In what follows, we will focus on a set of firm-level drivers that may explain differences across subgroups of firms.

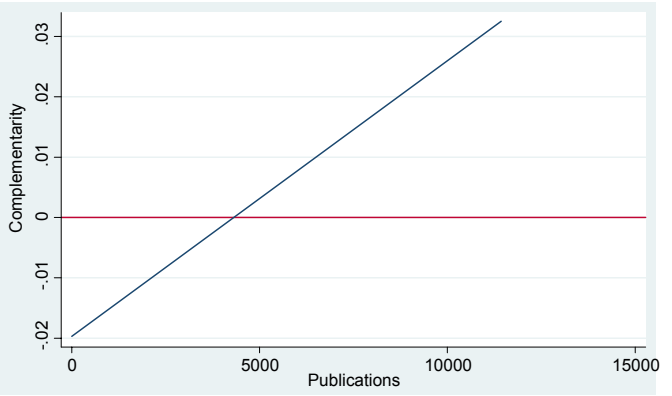
5.2 Firm-level Drivers of Complementarity

To identify the impact of potential drivers of complementarity, we first present a graphical analysis of the cross partial derivative (equation (5)) obtained using our benchmark GMM instrumental variable method with fixed effects. The objective of this analysis is to understand whether firms that perform better than others across the four different drivers experience a different level of complementarity among the two types of investments. Figure 1 reports the values of the degree of complementarity captured by the cross-partial over the range of our measures of absorptive capacity (scientific publications), economies of scope (number of therapeutic categories, or ATC), and licensing experience (stock of in-licensing investments) using our benchmark and full sample estimates presented in Table III, column 4.

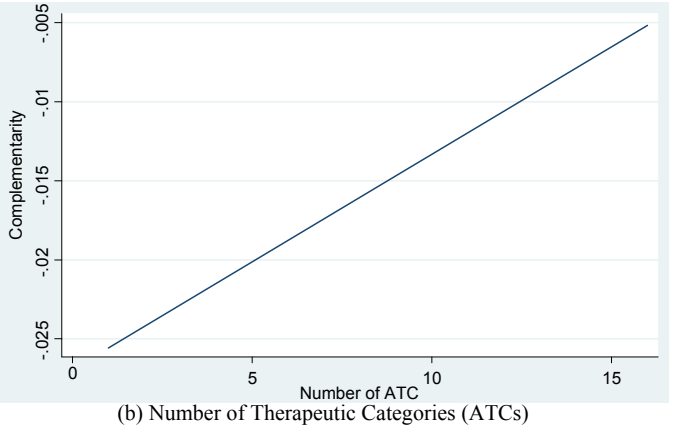
In all of the three graphs, the sign and magnitude of the joint effect vary with changes in the levels of the drivers. Overall, the cross-partial derivative (equation (5)) exhibits a positive trend in all cases, thus confirming that a higher level of complementarity is associated with higher levels of drivers.[15]

These findings confirm the complexity of the relationship between internal and external R&D investments; they also suggest that most previous studies on complementarity have not been able to ascertain whether a more composite relationship exists than can be revealed by the estimated average effect. A clear conclusion on whether two

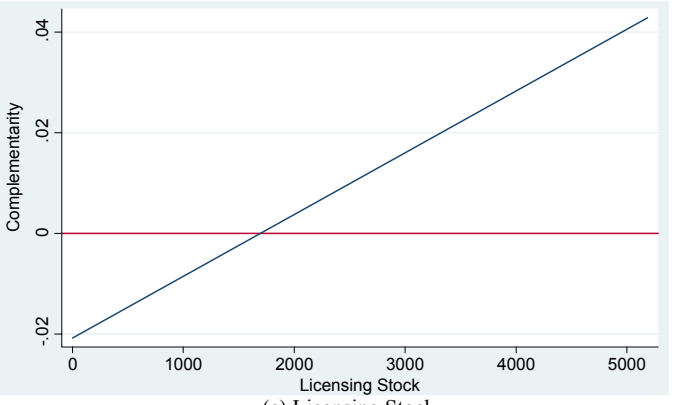
15. In what follows we provide examples of firms from the distribution plotted in Figure 1(a–c). That is, firms characterized by either small or low estimated cross-partials (complementarity indices) as a function of each driver. These include a variety of pharmaceutical and biotech companies of different sizes, and we take a cross-sectional perspective for descriptive purposes. Among firms in both the top quartiles of the scientific publications and cross-partial distributions (with positive complementarity) we find firms such as Eli Lilly, Bristol-Myers Squibb, and Genzyme. Among firms in both the bottom quartiles of the scientific publications and cross-partial distributions (with negative complementarity) we find Enzon, Mylan Laboratories, and AstraZeneca, among others. Among firms in both the top quartiles of therapeutic category (ATCs) and cross-partial distributions (with positive complementarity) we find firms such as Eli Lilly, GlaxoSmithKline, and Watson Pharmaceuticals. Among firms in both the bottom quartiles of the therapeutic category (ATCs) and cross-partial distributions (with negative complementarity) we find Repligen and Alcon, among others. Finally, among firms in both the top quartiles of licensing experience and cross-partial distributions (with positive complementarity) we find firms such as Elan, Bristol-Myers Squibb, and Abbott. Among firms in both the bottom quartiles of the licensing experience and cross-partial distributions (with negative complementarity) we find Enzon, Repligen, and Mylan Labs. Not surprisingly, firms with either low or high values of all drivers tend to have the lowest or the highest level of the cross-partial measuring complementarity, such as Enzon (with bottom 1% and negative value of the cross-partial) or Ely-Lilly (with top 1% and positive value of the cross-partial).



(a) Scientific Publications



(b) Number of Therapeutic Categories (ATCs)



(c) Licensing Stock

FIGURE 1. DISTRIBUTION OF COMPLEMENTARITY ACROSS DRIVERS (FULL SAMPLE ESTIMATION)

This figure provides graphical representations of the estimated cross-partial derivative $\frac{d^2n}{d\,R_i\,d\,R_e}$ *(Equation (5)) as a function of each examined driver of complementarity evaluated at each sample point, for example, for each firm in our sample. Estimates are based on results from our benchmark regression (Model 4, Table III).*

Table IV.
GMM Panel Estimates of the Translog Innovation Production Function (Split Samples)

	(1) Bottom 25% Publication	(2) Top 25% Publication	(3) Bottom 25% Number of ATCs	(4) Top 25% Number of ATCs	(5) Bottom 25% Licensing	(6) Top 25% Licensing
Ln R&D	0.468	1.428***	0.038	−0.163	−0.625	−0.406
	(0.361)	(0.115)	(0.227)	(0.287)	(0.443)	(1.528)
Ln licensing	−0.689***	0.761*	0.262	0.370	−0.723	2.951**
	(0.228)	(0.423)	(0.532)	(0.272)	(0.711)	(1.264)
(Ln R&D)^2	−0.093***	−0.071	0.040	−0.015	−0.002	−0.429***
	(0.025)	(0.051)	(0.065)	(0.045)	(0.057)	(0.165)
(Ln licensing)^2	0.095***	−0.022	−0.018	−0.098***	−0.061	−0.872***
	(0.023)	(0.034)	(0.073)	(0.028)	(0.138)	(0.290)
(Ln R&D)* (Ln licensing)	0.051	−0.012	−0.058	0.110**	0.352***	1.013***
	(0.072)	(0.086)	(0.078)	(0.053)	(0.116)	(0.348)
Publications	0.005	−0.221	0.094	−0.217***	−0.261*	−1.359***
	(0.077)	(0.169)	(0.084)	(0.045)	(0.141)	(0.353)
% Compound licensed—Phase I	−0.020	0.485	0.751	0.942**	−1.865**	0.879
	(0.132)	(0.380)	(0.627)	(0.449)	(0.915)	(1.206)
% Compound licensed—Phase II	0.268	−2.600***	0.191	−1.698***	−0.056	0.043
	(0.222)	(0.349)	(0.254)	(0.289)	(0.134)	(0.790)
% Compound licensed—Phase III	0.454*	−1.903***	−0.234**	−0.387	−0.465***	−1.599**
	(0.243)	(0.142)	(0.119)	(0.355)	(0.118)	(0.646)
Firm size	−0.072	0.003	0.007	0.001	−0.039	0.006
	(0.133)	(0.002)	(0.047)	(0.003)	(0.064)	(0.005)
Firm fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	95	124	177	159	112	111
Log-likelihood	−10.24	19.50	−5.274	−5.711	18.74	−36.50
Cluster	20	19	25	20	26	24
Over-identification test (p-value)	0.178	0.429	0.166	0.198	0.361	0.417
Endogenous RHS variables			*F*-statistic (excluded instruments from the Translog production function)			
Ln(RD)	19.38***	101.53***	29.11***	41.05***	5.81***	14.48***
Ln(licensing)	41.15***	36.45***	7.49***	5.07***	25.16***	8.94***
(Ln R&D)^2	13.75***	6.23***	5.19***	17.62***	6.55***	36.01***
(Ln licensing)^2	19.76***	33.77***	14.92***	18.91***	15.51***	10.42***
Ln(R&D)* Ln(licensing)	82.54***	8.67***	17.56***	12.65***	3.65**	25.77***

Notes: Standard errors clustered by firms in parentheses.
*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. The excluded instruments from the Translog innovation production function are the following: the natural logs, the square and cross-products of the natural logs of *Industry Promotion*, *Trademarks*, *Potential product market size, and Number of competitors*. ATC dummies vary over time because based on the primary ATC, which is defined in terms of annual sales.

Table V.
Tests on Mean Complementarity ($\frac{d^2n}{dR_i dR_e}$) Differences by Group of Firms

		Panel RE	Confidence Interval	Panel FE	Confidence Interval	Panel GMM FE	Confidence Interval
Scientific publications	≤25%	0.006		0.0039		0.0004	
	≥75%	−0.002		0.0006		0.018	
	Difference	−0.008	[−0.01, −0.006]	−0.003	[−0.004, −0.002]	0.018***	[0.009, 0.026]
Number of ATC	≤25%	−0.0105		−0.009		−0.008	
	≥75%	−0.004		0.004		0.0005	
	Difference	0.004***	[0.002, 0.01]	0.005***	[0.011, 0.017]	0.009***	[0.007, 0.01]
Licensing experience	≤25%	−0.035		0.041		−0.138	
	≥75%	−0.004		0.002		−0.026	
	Difference	0.029***	[0.019, 0.038]	−0.039	[−0.054, −0.024]	0.111***	[0.056, 0.167]

Notes: The table contains estimates of the average cross-partial $\frac{dn^2}{dR_i dR_e}$ (equation (5)) evaluated within subsamples of firms split based on observed values of the examined drivers. A positive mean difference suggests that firms above the top quartile of the distribution of the examined driver experience higher level of complementarity between internal and external R&D.
***, **, * indicate that the difference is > 0 at the 0.01, 0.05, and 0.1 confidence levels, respectively. We also report 95% confidence intervals for the difference between cross-partials.

activities are either complementary or substitute may be noninformative, because the joint effect changes across different ranges of value of key firm characteristics.

As a robustness measure, we present estimates in Table IV of the Translog production function (equation (2)) using our benchmark GMM method with firm-fixed effects within subsamples of firms characterized by either low (bottom 25%) or high (top 25%) levels of the distribution of the examined driver. We include the split-sample regression using the Panel Random Effect and Panel Fixed Effect methods in Table AII and AIII of the online Appendix, respectively.

For an analysis of the significance of the differences across groups, we present tests for mean complementarity differences across groups of firms defined using bottom and top quartiles of the distributions of the examined drivers for our split-sample estimations. A positive difference implies a higher degree of complementarity for the group of firms above the top quartile. These tests are shown in Table V.

Overall, our results confirm our expectations. We find that firms with strong scientific capabilities are characterized, on average, by a higher level of the cross-partial derivative capturing the degree of complementarity.[16] The results confirm that firms with broader experience across therapeutic areas are characterized, on average, by a stronger complementarity relationship between internal R&D and in-licensing. This finding suggests that these firms may be using knowledge developed in different fields additively in the innovative process, supporting the view espoused by Henderson and Cockburn (1996). Finally, results show that complementarity increases for firms that have a larger stock of prior licensing investments. This suggests that firms with greater levels of experience in licensing agreement formation facilitate the management and integration of external technologies.

6. Conclusion

Our goal has been to offer a deeper understanding of the exact nature of the relationship between internal R&D and in-licensing (external R&D). Although the extant literature

16. The result is supported in our benchmark estimation (Panel GMM fixed effect). We also report in Table AIV in the online Appendix the same tests using the stock of internal R&D and the stock of patents as alternative measures of absorptive capacity. The results are confirmed with our patent measure.

remains unclear about the relationship between these two strategies, our primary focus is to understand how the joint effect of two activities varies across several different drivers. Excluding the research by Cassiman and Veugelers (2006), there is a lack of empirical work examining the conditions under which internal R&D and in-licensing are either complements or substitutes.

We analyze possible determinants of this relationship by splitting our sample based on five potential drivers. Our mean tests confirm that complementarity appears to increase when associated with higher levels of select drivers. In other words, firms with higher absorptive capacity, those with alliance experience, and those that enjoy economies-of-scope are characterized by stronger complementarity. These results are confirmed by our graphical analysis and tests of hypotheses, which support a positive relation between complementarity and their drivers. Existing theories offer theoretical support of our results and provide insights for further theoretical work on the complementarity between innovative activities.

As is the case with all work, our study is subject to several limitations. First, we only analyze one dimension of innovative performance, the introduction of new drugs. In line with the contribution of Arora and Gambardella (1994), for example, one could claim that absorptive capacity, in particular a firm's scientific capability, will allow the technology buyer to be more discerning in the external technology that they select and will have a higher threshold value for each external R&D project. In other words, the mix of internal and external R&D may affect the expected value of an innovation, which we do not observe. To the extent that we are neglecting a potentially positive effect of the mix of internal and external R&D on the profitability of new drugs, our analysis can be considered as providing estimates of complementarity that are downward biased. This may contribute to explain why, on average, we do not find complementarity. In order to more fully analyze the marginal returns from internal and external R&D we would need data on the profitability associated with each drug, a task we leave for future work.

A second limitation relates to our identification strategy. Although the use of panel data and instrumental variable methods may increase our confidence on a causal interpretation of our results, we still achieve identification based on a specific set of economic assumptions, which could be violated and are not testable.

A third limitation of our study relates to our industry setting and the generalization of our results to other industries, because innovation factors are often determined by industrial dynamics. The R&D process in the pharmaceutical industry is characterized by long development cycles, high costs, and significant levels of uncertainty, which may affect the extent to which a firm relies on different innovative strategies. Industries that present a different innovative process might experience a different relationship. Although our results are not generalizable, our methodology can be replicated in different industry settings.

A fourth limitation lies in the definition and treatment of uncertainty associated with the drug development process. Recent research presents contrasting results about the possibility of success correlated to internally developed or externally acquired compounds. For example, Guedj (2005) shows that alliance projects are 21% more likely to move from Phase I to Phase II, whereas codeveloped compounds are less successful in later stages (Phase II, Phase III, and FDA approval) than internal projects. Conversely, Arora et al. (2009) suggest that asymmetric information and market imperfections increase costs, and the expected value of the licensed compound increases as a result. They show that the probability of success for a licensed compound is higher than for an internally developed one. We attempt to deal with the uncertainty related to

in-licensing investments by controlling for the percentage of in-licensed compounds in each phase. We also weight the firm's research pipeline by the average probability of success associated with each development stage to account for process development uncertainty.

A fifth limitation is methodological in nature. As already noted, although our Translog approximation to the true innovation production function is an improvement over the use other functional forms such the Cobb-Douglas, the Translog has some limitations and our results should be interpreted with caution.

Finally, while our results help us to understand the relation among innovative factors, we do not directly test whether there might be an optimal balance between R&D strategies, as suggested by other scholars. For example, Rothaermel et al. (2006) suggest that by performing some activities of the value chain internally and some externally, a firm is able to exploit external technology and adopt a flexible strategy to introduce new products. Knowing whether internal development and in-licensing are complements or substitutes might help build a feasible equilibrium between these two strategies. This would allow for a more complete understanding of the proposed outsourcing move by companies such as GlaxoSmithKline (Knowles and Higgins, 2011). Ultimately, this knowledge also allows for a deeper understanding of the feasibility of more radical views of the innovative process, such as the S&D model proposed by Morgan Stanley (2010).

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web site:

Online Appendix
Table AI. Panel GMM Estimation Using Subsets of Instruments
Table AII. Split Sample Regressions Using Random Effect Model
Table AIII. Split Sample Regressions Using Fixed Effect Model
Table AIV. GMM Panel Estimates of the Translog Innovation Production Function
Table SI. CES-Translog Estimation. Dependent Variable Ln(1 + pipeline)
Table SII. Specification Test Results Using CES-Translog Production Function

References

Acemoglu, D. and J. Linn, 2004, "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry," *The Quarterly Journal of Economics*, 119, 1049–1090.

Arora, A. and M. Ceccagnoli, 2006, "Patent Protection, Complementary Assets, and Firms' Incentives for Technology Licensing," *Management Science*, 52, 293–308.

———, ———, and W.M. Cohen, 2008, "R&D and the Patent Premium," *International Journal of Industrial Organization*, 26, 1153–1179.

———, A. Fosfuri, and A. Gambardella, 2001, *Markets for Technology: The Economics of Innovation and Corporate Strategy*, Cambridge, MA: MIT Press.

——— and A. Gambardella, 1994, "Evaluating Technological Information and Utilizing It: Scientific Knowledge, Technological Capability, and External Linkages in Biotechnology," *Journal of Economic Behavior and Organization*, 24, 91–114.

——— and ———, 2010, "Ideas for Rent: An Overview of Markets for Technology," *Industrial and Corporate Change*, 19, 775–803.

———, ———, L. Magazzini, and F. Pammolli, 2009, "A Breath of Fresh Air? Firm Type, Scale, Scope, and Selection Effects in Drug Development," *Management Science*, 55, 1638–1653.

Berndt, E.R. and L.R. Christensen, 1973, "The Translog Function and the Substitution of Equipment, Structures, and Labor in U.S. Manufacturing 1929–68," *Journal of Econometrics*, 1, 81–113.

Bresnahan, T. and A. Gambardella, 1998, "The Division of Inventive Labor and the Extent of the Market," in E. Helpman, ed. *General Purpose Technologies and Economic Growth*, Cambridge, MA: MIT.

Cassiman, B. and R. Veugelers, 2006, "In Search of Complementarity in Innovation Strategy: Internal R&D and External Knowledge Acquisition," *Management Science*, 52, 68–82.

Chan, T., J.A. Nickerson, and H. Owan, 2007, "Strategic Management of R&D Pipelines with Cospecialized Investments and Technology Markets," *Management Science*, 53, 667–682.

Cohen, W.M. and D.A. Levinthal, 1989, "Innovation and Learning: The Two Faces of R&D," *Economic Journal*, 99, 569–596.

Danzon, P.M., S. Nicholson, and N.S. Pereira, 2005, "Productivity in Pharmaceutical–Biotechnology R&D: The Role of Experience and Alliances," *Journal of Health Economics*, 24, 317–339.

———, A. Epstein, and S. Nicholson, 2007, "Mergers and Acquisitions in the Pharmaceutical and Biotech Industries," *Managerial and Decision Economics*, 28, 307–328.

Forman, C., A. Goldfarb, and S. Greenstein, 2008, "Understanding the Inputs into Innovation: Do Cities Substitute for Internal Firm Resources?," *Journal of Economics & Management Strategy*, 17, 295–316.

Fosfuri, A., M.S. Giarratana, and A. Luzzi, 2008, "The Penguin Has Entered the Building: The Commercialization of Open Source Software Products," *Organization Science*, 19, 292–305.

Glaeser, E.L., D.I. Laibson, J.A. Scheinkman, and C.L. Soutter, 2000, "Measuring Trust," *The Quarterly Journal of Economics*, 115, 811–846.

Grabowski, H., 2002, "Patents, Innovation and Access to New Pharmaceuticals," *Journal of International Economic Law*, 5, 849–860.

Greene, W.H., 2012, *Econometric Analysis*, seventh edition, Upper Saddle River, NJ: Prentice Hall.

Guedj, I., 2005,"Ownership vs. Contract: How Vertical Integration Affects Investment Decisions in Pharmaceutical R&D," *McCombs Research Paper Series No. FIN-01–06. SSRN*. Available at: http://ssrn.com/abstract=677371 (accessed June 6, 2012).

Hall, B.H., 1993, "Industrial Research During the 1980s: Did the Rate of Return Fall?," *Brookings Papers on Economic Activity: Microeconomics*, 2, 289–343.

Hayashi, F., 2000, "Econometrics. 2000," *Princeton University Press. Section*, 1, 60–69.

Henderson, R. and I. Cockburn, 1996, "Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery," *RAND Journal of Economics*, 27, 32–59.

Higgins, M.J. and D. Rodriguez, 2006, "The Outsourcing of R&D through Acquisitions in the Pharmaceutical Industry," *Journal of Financial Economics*, 80, 351–383.

Knowles, S.M. and M. Higgins, 2011, "Vertical Disintegration in the Pharmaceutical Industry and the Role of Ip," *Intellectual Asset Management*, 45, 10–15.

Laursen, K. and A. Salter, 2006, "Open for Innovation: The Role of Openness in Explaining Innovation Performance among U.K. Manufacturing Firms," *Strategic Management Journal*, 27, 131–150.

Lowe, J. and P. Taylor, 1998, "R&D and Technology Purchase through Licence Agreements: Complementary Strategies and Complementary Assets," *R&D Management*, 28, 263–278.

Malerba, F. and L. Orsenigo, 2000, "Knowledge, Innovative Activities and Industrial Evolution," *Industrial and Corporate Change*, 9, 289–314.

Milgrom, P. and J. Roberts, 1990, "The Economics of Modern Manufacturing: Technology, Strategy, and Organization," *American Economic Review*, 80, 511–528.

Morgan Stanley Research Europe, 2010, "Pharmaceuticals: Exit Research and Create Value," Morgan Stanley, New York.

Panzar, J.C. and R.D. Willig, 1981, "Economies of Scope," *American Economic Review*, 71, 268–272.

Pisano, G.P., 1990, "The R&D Boundaries of the Firm: An Empirical Analysis," *Administrative Science Quarterly*, 35, 153–176.

Pollak, R.A., R.C. Sickles, and T.J. Wales, 1984, "The Ces-Translog: Specification and Estimation of a New Cost Function," *The Review of Economics and Statistics*, 66, 602–607.

Rothaermel, F.T., M.A. Hitt, and L.A. Jobe, 2006, "Balancing Vertical Integration and Strategic Outsourcing: Effects on Product Portfolio, Product Success, and Firm Performance," *Strategic Management Journal*, 27, 1033–1056.

Scherer, F., 2007, "Pharmaceutical Innovation," *AEI-Brookings Center Working Paper*.

Teece, D.J., 1986, "Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy," *Research Policy*, 15, 285–305.

Tollman, P., Y. Morieux, J.K. Murphy, and U. Schulze, 2011, "Identifying R&D Outliers," *Nature Reviews Drug Discovery*, 10, 653–654.

Tsai, K.-H. and J.-C. Wang, 2008, "External Technology Acquisition and Firm Performance: A Longitudinal Study," *Journal of Business Venturing*, 23, 91–112.

Vega-Jurado, J., A. Gutiérrez-Gracia, and I. Fernández-de-Lucio, 2009, "Does External Knowledge Sourcing Matter for Innovation? Evidence from the Spanish Manufacturing Industry," *Industrial and Corporate Change*, 18, 637–670.

Vives, X., 2008, "Innovation and Competitive Pressure," *The Journal of Industrial Economics*, 56, 419–469.

Zollo, M. and S.G. Winter, 2002, "Deliberate Learning and the Evolution of Dynamic Capabilities," *Organization Science*, 13, 339–351.