

NBER WORKING PAPER SERIES

BEHIND THE SCENES:
SOURCES OF COMPLEMENTARITY IN R&D

Marco Ceccagnoli
Matthew J. Higgins
Vincenzo Palermo

Working Paper 18795
<http://www.nber.org/papers/w18795>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2013

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, the 2011 CCC doctoral consortium at MIT, the 2011 Academy of Management meeting, the 2011 DRUID conference, the 2012 Trans-Atlantic Doctoral Conference at the London Business School for helpful comments and suggestions. We are indebted to Deloitte ReCap for access to their data. We also thank Alexandra Kondo and 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, any of its affiliated or subsidiary entities, or the National Bureau of Economic Research. 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.

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 They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2013 by Marco Ceccagnoli, Matthew J. Higgins, and Vincenzo Palermo. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Behind the Scenes: Sources of Complementarity in R&D
Marco Ceccagnoli, Matthew J. Higgins, and Vincenzo Palermo
NBER Working Paper No. 18795
February 2013
JEL No. L24,L65,O31,O32

ABSTRACT

Even though management consultants increasingly recommend that in-house research be outsourced, little is known about the conditions favoring substitution or complementarity between internal R&D and external technology acquisition. In this paper, we attempt to provide a deeper understanding of the firm-level drivers of complementarity between these two types of investments through the structural estimation of a flexible innovation production function, such as the translog. Our empirical analysis is based on a unique panel dataset on the R&D and in-licensing expenditures of 94 global pharmaceutical firms active in drug development between 1997 and 2005. Our results suggest that internal R&D and in-licensing in the pharmaceutical industry were neither complements nor substitutes during the study period. However, we find that the degree of complementarity is enhanced for firms with stronger absorptive capacity, economies of scope, and past licensing experience.

Marco Ceccagnoli
Scheller College of Business
Georgia Institute of Technology
800 West Peachtree Street, NW
Atlanta, GA 30308
marco.ceccagnoli@scheller.gatech.edu

Vincenzo Palermo
Scheller College of Business
Georgia Institute of Technology
800 West Peachtree Street
Atlanta, GA 30308
Vincenzo.Palermo@scheller.gatech.edu

Matthew J. Higgins
Scheller College of Business
Georgia Institute of Technology
800 West Peachtree Street
Atlanta, GA 30308
and NBER
matt.higgins@scheller.gatech.edu

1. Introduction

Markets for technology have been extensively studied (Arora and Gambardella, 1990, 1994a); 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, 1990). 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 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. While 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 (GSK) 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 non-integrated 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 adopt a two-step empirical strategy. In the first step, we estimate the coefficients of a flexible CES-Translog innovation production function (Pollak et al., 1984) to find the most appropriate functional form to use in our context. We then provide structural estimates of the degree of complementarity or substitutability between these two types of R&D investments that vary across firms and time.

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 global 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, since 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 Fosfuri, 2003; 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. While 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; Cassiman and Veugelers, 2006; Tsai and Wang, 2008; Png, 2012). 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). The study of Tsai and Wang (2008) on Taiwanese electronics manufacturing demonstrates that external technology acquisition does not contribute to firm performance *per se*, but does show that external acquisition of technology has a positive effect on performance when interacting 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 allows—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 literature review suggests that the 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 (1994b) 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 re-adopted 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. While 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 above. 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; Hoang and Rothaermel, 2010). 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,” in order to allow external technologies from their various partnerships to become more accessible to internal projects.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 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

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

3.1. Step 1: CES-Translog specification and functional form tests

We assume that each firm 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):

$$(1) \quad n = f(R_i, R_e, S)$$

Hereafter, the firm and time subscripts are omitted for simplicity. We start by adopting a CES-Translog specification, a flexible, functional form that nests the Cobb-Douglas, CES, and Translog specifications. Previous work using a CES-Translog specification include Pollak et al. (1984) and Dewan and Min (1997). The former provide estimates of a CES-Translog cost function and find that it fits significantly better multiple datasets covering different industries and countries than had been previously used in the literature. The latter builds upon Pollak et al. (1984) to directly estimate the CES-Translog as a production function to analyze the effect of IT and non-IT capital on productivity in the IT industry.

There are two main advantages of using a CES-Translog specification. First, it is a flexible, functional form that is compatible with a wider range of substitution possibilities than CES or Translog. Second, we can exploit its nested properties to find the best functional form that describes the innovative process without losing efficiency in terms of likelihood (Pollak et al., 1984) and without imposing *a priori* restrictions on our model. Overall, this methodology helps improve upon existing research, because estimating the complementarity relationship between internal and external R&D using restrictive production models may cause specification errors and yield biased econometric estimates.

We define our CES-Translog production function as:

$$(2) \quad \log n = S + \frac{1}{\rho} \log(\alpha_i R_i^{-\rho} + \alpha_e R_e^{-\rho}) + \beta_i (\log R_i)^2 + \beta_e (\log R_e)^2 + \gamma_{ie} \log R_i \log R_e + u,$$

where R_i and R_e represent internal R&D expenditure and in-licensing investment (external R&D), $\alpha_i + \alpha_e = 1$, and u is a random error term representing the unobserved drivers of the internal and external R&D investments. 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 firms or universities. The additive linear term is equivalent to a classic CES specification, where ρ represents the elasticity of substitution between R_i and R_e . The β s and γ coefficients represent, respectively, the quadratic impact and the cross-effect of R&D investments on the production of innovations.²

According to Pollak et al. (1984), researchers can exploit the nested properties of the CES-Translog to test for other functional forms, and to reduce model complexity without reducing its estimation efficiency. These tests offer more flexibility to researches in modeling innovation functions, because they do not require that prior assumptions on data behavior be made. The Cobb-Douglas, CES, and Translog forms are all special cases of the CES-Translog. When all quadratic terms are equal to zero, we obtain a CES function. The Cobb-Douglas is obtained when all quadratic terms are equal to zero and ρ tends to zero. The Translog specification can be found when ρ approaches 0 and all the other parameters are different from zero. These nested properties enable testing for model specification using conventional test procedures (Pollak et al., 1984). Table 1 summarizes these three specification tests. A rejection of all the specification tests presented in Table 1 would lead to the adoption of the CES-Translog innovation production function. However, a simpler, but still efficient functional form can be adopted if one of the tests is rejected.

<Insert Table 1 here>

3.2.Step 2: Estimation of the degree of complementarity

² Our definition of the degree of complementarity/substitutability is based on the cross partial derivative of the production function. This is different from the elasticity of substitution, which is defined as the percentage change in factor proportions due to a change in the marginal rate of technical substitution ((Hicks, 1932)

After testing the coefficients of the CES-Translog, we are able to choose the most efficient functional form specification to be used in our model. Our empirical findings, discussed below in Section 5.1, will show that the Translog production function better fits our data. Therefore, we will only compute and focus on the degree of complementarity or substitutability for the Translog production function in the sections that follow.

The Translog specification is defined as follows:

$$(3) \quad n = R_i^{\alpha_i} R_e^{\alpha_e} e^{S + \beta_i (\log R_i)^2 + \beta_e (\log R_e)^2 + \gamma_{ie} \log R_i \log R_e + u},$$

where S is the exogenous component of the production function, u is the error component, and R_i and R_e represent internal and external research, respectively. We estimate the Translog by taking logarithms of both sides of (3), which allow us to employ linear estimation techniques.

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

$$(4) \quad \frac{dn}{dR_i} = \frac{n}{R_i} (\alpha_i + 2\beta_i \log R_i + \gamma_{ie} \log R_e) = \frac{n}{R_i} Z_i$$

$$(5) \quad \frac{dn}{dR_e} = \frac{n}{R_e} (\alpha_e + 2\beta_e \log R_e + \gamma_{ie} \log R_i) = \frac{n}{R_e} Z_e$$

We then estimate the degree of complementarity or substitutability using the following cross-partial derivative:

$$(6) \quad \frac{d^2n}{dR_i dR_e} = \left[\frac{n}{R_e} (\alpha_e + 2\beta_e \log R_e + \gamma_{ie} \log R_i) \frac{Z_i}{R_i} + \frac{n \gamma_{ie}}{R_i R_e} \right] = \frac{n}{R_i R_e} (Z_i Z_e + \gamma_{ie}) = \frac{n}{R_i R_e} \tilde{Z},$$

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 $\frac{d^2n}{dR_i dR_e}$ 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 complementary or

substitutable. 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 (3). We then evaluate $\frac{d^2 n}{dR_i dR_e}$ 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.

3.3. Empirical strategy

Our estimation procedure involves two steps. First, we identify whether the innovation production function is better represented by a Cobb-Douglas, CES, Translog, or CES-Translog by estimating and testing the coefficients of Equation (2). Second, we select the best functional form to estimate the degree of complementarity/substitutability. We are able to estimate all model parameters regardless of the functional form of the production function. Therefore, all the equations are identified. Once the innovation production function is estimated, we can compute the sign and magnitude of the cross-partial derivative $\frac{d^2 n}{dR_i dR_e}$.

Notice that $\gamma_{ie} \equiv \frac{d^2 \log n}{d \log R_i d \log R_e}$ in (3)-(6) represents a percentage change in the elasticity of internal R&D for a percentage change in licensing or vice versa. While this is a more easily interpretable notion of complementarity, an evaluation of the sign of the cross-partial derivative (6) suggests that its sign is not determined by the sign of γ_{ie} .

The production functions presented in the previous sections can be used in the context of a profit maximization model with endogenous internal and external R&D investment levels (available from the authors upon request). Such models generate exclusion restrictions which imply that variables affecting the optimal level of internal R&D and licensing (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 R&D and licensing. 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 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 dataset 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 includes 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 and the novelty of the compound.

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, up-front payments, and milestones. Finally, from IMS MIDAS™ we obtained product-level promotion expenditures. All financial variables are in year 2000 constant US dollars. Descriptive statistics are provided in Table 2 and correlations are presented in Table 3.

<Insert Table 2 and Table 3 here>

Our final sample consists of 94 global 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 U.K. 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 firms 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.³

4.1. Dependent variables

Product pipeline. Our dependent variable is the firm-year product pipeline, which represents a firm's innovative output. The importance of studying a firm 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 those being developed for each firm in our sample.⁴ To account for development uncertainty, compounds are weighted by average probabilities of successfully reaching 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

Internal R&D investments. We compute internal R&D investments using data from Compustat and Deloitte ReCap. R&D data from Compustat includes expenditures in R&D

³ 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.

⁴ 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.

that could be performed internally or externally.⁵ 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).

In-licensing investment (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.⁶ 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.⁷ Because the stock of licensing expenditures also capture a firm's licensing experience, we will 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.

4.3. Instrumental variables for internal and external R&D

⁵ 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.”

⁶ 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.

⁷ 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.

As noted above, 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 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.⁸ We therefore use variables that should affect the profitability of marketed drugs, such as potential size of the market, drug novelty, number of competitors, and the strength of a firm's complementary assets. As shown in the empirical results section, the above instruments appear to have sufficient power and seem to be uncorrelated with the econometric error term, as indicated by the tests for instrument validity.

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 drug novelty as another potential instrument. This data is made available by Pharmaprojects, which contains independent ratings about the novelty of compounds. Each compound's novelty is categorized using a discrete range between 1 and 6, where the value 6 represents the most innovative drugs. Consistent with recent work, our measure of drug novelty is based on drugs with the highest novelty rank, corresponding to "leading compounds" (Grabowski and Wang, 2006; DiMasi and Faden, 2011; Azoulay and

⁸ The formal optimization model is available from the authors upon request.

Fishman, 2013). We then use the proportion of novel drugs in the pipeline for each firm-year of each firm as an additional instrumental variable.

Third, we use the number of competitors to proxy for the incentive to be innovative and productive. While 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. The data was collected from IMS MIDAS™.

Finally, ownership and strength of downstream complementary assets is an important driver of the appropriability of returns from innovation (Teece, 1986). We use two variables to proxy for a firm's strength of complementary assets. First, we employ a firm's detailing expenditures, obtained from IMS MIDAS™, to capture that firm's marketing capability. Detailing is defined as promotion activities directed toward physicians and hospitals, journal advertising, and direct-mail. We also use the stock of a firm's trademarks to proxy for that firm's brand capital (Fosfuri et al., 2008). These two instruments are associated with both internal and external R&D investments, since they enhance the appropriability of both types of investments. We collect data on active trademarks from the USPTO and use them to build a stock variable.

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 is 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 expenditures (which includes applied research and development 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, as measures of a firm’s type of R&D we use the focal firm’s 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 (1994b) and utilize two alternative measures: 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 that has already gone through part of the earliest and more uncertain stages of the development process.⁹

Finally, we control for industry, firm, location, and year of unobserved fixed effects. 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 area with the highest level of annual sales, then we include a set of dummy variables that would 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 dummies and controls for 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 dummies and geographic location dummies (North America, Europe, and other).

5. Results

5.1. First step: Functional form tests

Our estimation procedure starts by estimating the coefficients of a CES-Translog production function (Equation 2). The tests are summarized in Table 1. The advantage of adopting a flexible specification in the first step is due to its nested properties. Equation (2) allows us to test whether the production function can be simplified by using a Cobb-Douglas, CES, or Translog function. Our regression specification tests are reported in Table 4. In this first set of analyses, we are not interested in the marginal effect of our independent variables but rather focus only on the specification tests described in Table 1. Marginal effects and the

⁹ 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.

degree of complementarity or substitutability, if any, are the focus of the second step of our empirical estimation procedure, discussed below.

<Insert Table 4 here>

We report the results for different models in Table 4. Model (1) is estimated using OLS Fixed Effect. It includes our main variables (internal R&D and licensing expenditures) and our full set of controls, including firm fixed-effects. Model (2) is estimated using GMM and incorporates the main variables and a full set of controls.^{10,11} The instrumental variables pass the validity tests, as discussed in Table 4.

Our results, which are robust across the estimated models, indicate that ρ is not significantly different from zero. As a result, we can adopt a Translog specification for our production function, as defined by Equation (4). Moreover, we clearly reject the possible use of both a Cobb-Douglas and CES specification, because the related tests specified in Table 1 are significant. While ρ is not significantly different from zero, all the coefficients on the quadratic terms (β_i , β_e , and γ_{ie} , respectively) are jointly different from zero.

5.2.Second step: Estimating the degree of complementarity/substitutability using the Translog production function.

After identifying Equation (4) as the appropriate innovation production function, we focus on estimating 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 elasticities and the

¹⁰ The nonlinearity of Equation 4 does not allow us to eliminate the firm fixed effects using first-differences, nor by transforming the data to within-firm deviations. As a result, we control for unobserved firm heterogeneity using a set of firm-specific dummy variables.

¹¹ The GMM model with all controls (2) and all firm fixed-effects, however, did not converge. Thus, it is not reported. Details are available from the authors upon request.

degree of complementarity by taking the derivative with respect to the logarithm of internal R&D and in-licensing (external R&D).

The estimates of the Translog are reported in Table 5. 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 6) associated with the models of Table 5 are reported in Table 6. The cross-partials are evaluated at the mean and median of the variables included in Equation (6).

<Insert Tables 5-6 here>

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 6 are not significantly different from zero across estimation methods.

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

The results presented in Table 6 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. Furthermore, our empirical approach offers a new method to estimate complementarity without imposing methodological restrictions on the estimation, which could bias the results. 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.

5.3. 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 $\frac{d^2n}{dR_i dR_e}$ (Equation 6) 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, stock of internal R&D, patents), economies of scope (number of therapeutic categories, or ATC), and licensing experience (stock of in-licensing investments).

<Insert Figure 1 here>

In all of the five graphs, the sign and magnitude of the joint effect vary with changes in the levels of the drivers. Overall, $\frac{d^2n}{dR_i dR_e}$ exhibits a positive trend in all cases, thus confirming that a higher level of complementarity is associated with higher levels of drivers. 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 activities are either

complementary or substitute may be non-informative, since the joint effect changes across different ranges of value of key firm characteristics.

As a robustness measure, we present estimates in Table 7a and 7b of the Translog production function (Equation 4) using our benchmark GMM method with firm-fixed effects within sub-samples of firms characterized by either low (bottom 25%) or high (top 25%) levels of the distribution of the examined driver.

<Insert Tables 7a-b here>

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 both full and split-sample estimations. The resulting difference's positive value implies a higher degree of complementarity for the group of firms above the top quartile. These tests are shown in Table 8.

<Insert Table 8 here>

Overall, our results confirm our expectations. We find that firms with highly cumulative levels of scientific publications, internal R&D, or patents are characterized, on average, by a higher level of the cross-partial derivative capturing the degree of complementarity. The results confirm that firms with broader experiences 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, which would support Henderson and Cockburn (1996) view. Finally, results indicate that complementarity increases for firms that have a larger stock of prior licensing deals, which possibly indicates that higher levels of experience in licensing agreement formation facilitate the management and integration of external technologies (Hoang and Rothaermel, 2010).

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). While the extant literature 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 the selected 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 drivers. Existing theories offer theoretical support of our results and provide insights for further theoretical work on the complementarity between innovative activities. At the same time, we provide a methodological contribution, since our framework can be used for a more rigorous understanding of the industry and firm characteristics that affect the relationship between internal and external innovative activities.

One limitation of this research comes from the fact that we only analyze one dimension of innovative performance, the introduction of new drugs. In line with the contribution of Arora and Gambardella (1994b), 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 of our study relates to our industry setting and the generalization of our results to other industries, since 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 third 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, while co-developed 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.

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 search and development model proposed by Morgan Stanley (2010).

References

- Acemoglu D., Linn J., 2004, "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry," *The Quarterly Journal of Economics*, **119**(3), 1049-1090.
- Arora A., Ceccagnoli M., 2006, "Patent Protection, Complementary Assets, and Firms' Incentives for Technology Licensing," *Management Science*, **52**(2), 293-308.
- Arora A., Fosfuri A., 2003, "Licensing the Market for Technology," *Journal of Economic Behavior and Organization*, **52**(2), 277-295.
- Arora A., Fosfuri A., Gambardella A., 2001, *Markets for Technology: The Economics of Innovation and Corporate Strategy*. MIT Press: Cambridge, MA.
- Arora A., Gambardella A., 1990, "Complementarity and External Linkages: The Strategies of the Large Firms in Biotechnology," *The Journal of Industrial Economics*, **38**(4), 361-379.
- Arora A., Gambardella A., 1994a, "The Changing Technology of Technological Change: General and Abstract Knowledge and the Division of Innovative Labour," *Research Policy*, **23**(5), 523-532.
- Arora A., Gambardella A., 1994b, "Evaluating Technological Information and Utilizing It: Scientific Knowledge, Technological Capability, and External Linkages in Biotechnology," *Journal of Economic Behavior and Organization*, **24**(1), 91-114.
- Arora A., Gambardella A., 2010, "Ideas for Rent: An Overview of Markets for Technology," *Industrial and Corporate Change*, **19**(3), 775-803.
- Arora A., Gambardella A., Magazzini L., Pammolli F., 2009, "A Breath of Fresh Air? Firm Type, Scale, Scope, and Selection Effects in Drug Development," *Management Science*, **55**(10), 1638-1653.
- Azoulay P., Fishman A., 2013, "Drugs, \$\$, and Drug Development: The Rise of for-Profit Experimental Medicine." In *Factors Affecting the Location of Biopharmaceutical Activities*. Cockburn IM, Slaughter MJ (eds.), University of Chicago Press. Forthcoming: Chicago, IL.
- Bresnahan T., Gambardella A., 1998, "The Division of Inventive Labor and the Extent of the Market." In *General Purpose Technologies and Economic Growth*. Helpman E (ed.), MIT: Cambridge, MA.
- Cassiman B., Veugelers R., 2006, "In Search of Complementarity in Innovation Strategy: Internal R&D and External Knowledge Acquisition," *Management Science*, **52**(1), 68-82.

Chan T., Nickerson J. A., Owan H., 2007, "Strategic Management of R&D Pipelines with Cospecialized Investments and Technology Markets," *Management Science*, **53**(4), 667-682.

Cohen W. M., Levinthal D. A., 1989, "Innovation and Learning: The Two Faces of R&D," *Economic Journal*, **99**(397), 569-596.

Danzon P. M., Epstein A., Nicholson S., 2007, "Mergers and Acquisitions in the Pharmaceutical and Biotech Industries," *Managerial and Decision Economics*, **28**(4-5), 307-328.

Danzon P. M., Nicholson S., Pereira N. S., 2005, "Productivity in Pharmaceutical-Biotechnology R&D: The Role of Experience and Alliances," *Journal of Health Economics*, **24**(2), 317-339.

Dewan S., Min C.-k., 1997, "The Substitution of Information Technology for Other Factors of Production: A Firm Level Analysis," *Management Science*, **43**(12), 1660-1675.

DiMasi J. A., Faden L. B., 2011, "Competitiveness in Follow-on Drug R&D: A Race or Imitation?," *Nature Reviews Drug Discovery*, **10**(1), 23-27.

Forman C., Goldfarb A., Greenstein S., 2008, "Understanding the Inputs into Innovation: Do Cities Substitute for Internal Firm Resources?," *Journal of Economics & Management Strategy*, **17**(2), 295-316.

Fosfuri A., Giarratana M. S., Luzzi A., 2008, "The Penguin Has Entered the Building: The Commercialization of Open Source Software Products," *Organization Science*, **19**(2), 292-305.

Grabowski H., 2002, "Patents, Innovation and Access to New Pharmaceuticals," *Journal of International Economic Law*, **5**(4), 849-860.

Grabowski H. G., Wang R. Y., 2006, "The Quantity and Quality of Worldwide New Drug Introductions, 1982-2003," *Health Affairs*, **25**(2), 452-460.

Guedj I., 2005, "Ownership Vs. Contract: How Vertical Integration Affects Investment Decisions in Pharmaceutical R&D." In *McCombs Research Paper Series No. FIN-01-06*. SSRN: <http://ssrn.com/abstract=677371>

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

Henderson R., Cockburn I., 1996, "Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery," *RAND Journal of Economics*, **27**(1), 32-59.

Hicks J., 1932, "Marginal Productivity and the Principle of Variation," *Economica*, **35**79-88.

Higgins M. J., Rodriguez D., 2006, "The Outsourcing of R&D through Acquisitions in the Pharmaceutical Industry," *Journal of Financial Economics*, **80**(2), 351-383.

Hoang H., Rothaermel F. T., 2010, "Leveraging Internal and External Experience: Exploration, Exploitation, and R&D Project Performance," *Strategic Management Journal*, **31**(7), 734-758.

Knowles S. M., Higgins M., 2011, "Vertical Disintegration in the Pharmaceutical Industry and the Role of Ip," *Intellectual Asset Management*, **4**510-15.

Laursen K., Salter A., 2006, "Open for Innovation: The Role of Openness in Explaining Innovation Performance among U.K. Manufacturing Firms," *Strategic Management Journal*, **27**(2), 131-150.

Lowe J., Taylor P., 1998, "R&D and Technology Purchase through Licence Agreements: Complementary Strategies and Complementary Assets," *R&D Management*, **28**(4), 263-278.

Malerba F., Orsenigo L., 2000, "Knowledge, Innovative Activities and Industrial Evolution," *Industrial and Corporate Change*, **9**(2), 289-314.

Milgrom P., Roberts J., 1990, "The Economics of Modern Manufacturing: Technology, Strategy, and Organization," *American Economic Review*, **80**(3), 511-528.

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

Panzar J. C., Willig R. D., 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**(1), 153-176.

Png I. P. L., 2012, "Law and Innovation: Evidence from State Trade Secrets Laws." SSRN eLibrary. <http://ssrn.com/abstract=1755284>.

Pollak R. A., Sickles R. C., Wales T. J., 1984, "The Ces-Translog: Specification and Estimation of a New Cost Function," *The Review of Economics and Statistics*, **66**(4), 602-607.

Rothaermel F. T., Hitt M. A., Jobe L. A., 2006, "Balancing Vertical Integration and Strategic Outsourcing: Effects on Product Portfolio, Product Success, and Firm Performance," *Strategic Management Journal*, **27**(11), 1033-1056.

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

Teece D. J., 1986, "Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy," *Research Policy*, **15**(6), 285-305.

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

Tsai K.-H., Wang J.-C., 2008, "External Technology Acquisition and Firm Performance: A Longitudinal Study," *Journal of Business Venturing*, **23**(1), 91-112.

Vega-Jurado J., Gutiérrez-Gracia A., Fernández-de-Lucio I., 2009, "Does External Knowledge Sourcing Matter for Innovation? Evidence from the Spanish Manufacturing Industry," *Industrial and Corporate Change*, **18**(4), 637-670.

Vives X., 2008, "Innovation and Competitive Pressure," *The Journal of Industrial Economics*, **56**(3), 419-469.

Zollo M., Winter S. G., 2002, "Deliberate Learning and the Evolution of Dynamic Capabilities," *Organization Science*, **13**(3), 339-351.

Table 1. Functional form tests

| Functional form | Coefficients test |
|------------------------|---|
| Cobb-Douglas | $\rho = 0; \beta_i = 0; \beta_e = 0; \gamma_{ie} = 0$ |
| CES | $\beta_i = 0; \beta_e = 0; \gamma_{ie} = 0$ |
| Translog | $\rho = 0$ |

Coefficients from the CES-Translog production function (Equation 2).

Table 2. Descriptive Statistics

| Variable | Mean | Std. Dev. | Min | Max |
|---|-------------|------------------|------------|------------|
| Product pipeline | 1.764 | 1.184 | 0 | 5.029 |
| In-Licensing (deflated, Mil. \$) | 239.443 | 633.850 | 0 | 5184.333 |
| Internal R&D (deflated, Mil. \$) | 1345.112 | 3195.205 | 0.473 | 28756.440 |
| Detailing stock (deflated, Thousands \$) | 7042.601 | 19559.700 | 0 | 173521.400 |
| Trademark stock | 11.590 | 34.003 | 0 | 426 |
| Expected market size (deflated, Thousands \$) | 2158.068 | 1556.527 | 0 | 10217.840 |
| Competitors | 1308.882 | 578.575 | 201 | 2626 |
| Drug novelty | 0.154 | 0.185 | 0 | 1 |
| Sales (deflated, Mil. \$) | 4510.527 | 10029.110 | 0 | 67674.560 |
| Firm size (hundreds) | 13.841 | 26.717 | 0.001 | 122 |
| Scientific References | 3.689 | 3.785 | 0 | 36 |
| 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

Table 3. Correlation Table

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|--|--------|-------|-------|-------|-------|--------|--------|--------|-------|-------|-------|-------|-------|-------|----|
| 1. Product pipeline | 1 | | | | | | | | | | | | | | |
| 2. In-Licensing (deflated, Mil. \$) | 0.492 | 1 | | | | | | | | | | | | | |
| 3. Internal R&D (deflated, Mil. \$) | 0.629 | 0.733 | 1 | | | | | | | | | | | | |
| 4. Detailing stock (deflated, Thousands \$) | 0.642 | 0.783 | 0.937 | 1 | | | | | | | | | | | |
| 5. Trademark stock | 0.338 | 0.549 | 0.668 | 0.597 | 1 | | | | | | | | | | |
| 6. Expected market size (deflated, Thousands \$) | 0.109 | 0.090 | 0.156 | 0.110 | 0.096 | 1 | | | | | | | | | |
| 7. Competitors | 0.054 | 0.163 | 0.167 | 0.169 | 0.092 | 0.179 | 1 | | | | | | | | |
| 8. Drug novelty | -0.069 | 0.039 | 0.023 | 0.020 | 0.017 | 0.264 | -0.049 | 1 | | | | | | | |
| 9. Sales (deflated, Mil. \$) | 0.543 | 0.453 | 0.738 | 0.622 | 0.493 | 0.249 | 0.166 | -0.045 | 1 | | | | | | |
| 10. Firm size (hundreds) | 0.640 | 0.493 | 0.809 | 0.680 | 0.542 | 0.244 | 0.142 | -0.036 | 0.902 | 1 | | | | | |
| 11. Scientific References | 0.301 | 0.217 | 0.239 | 0.214 | 0.177 | -0.041 | -0.020 | -0.044 | 0.177 | 0.251 | 1 | | | | |
| 12. Number of ATCs | 0.649 | 0.398 | 0.499 | 0.513 | 0.322 | 0.195 | 0.177 | -0.135 | 0.580 | 0.638 | 0.218 | 1 | | | |
| 13. % licensed compound (Phase 1) | 0.157 | 0.137 | 0.129 | 0.103 | 0.121 | 0.162 | 0.154 | 0.004 | 0.085 | 0.119 | 0.084 | 0.100 | 1 | | |
| 14. % licensed compound (Phase 2) | 0.072 | 0.086 | 0.097 | 0.068 | 0.049 | -0.010 | 0.127 | 0.049 | 0.100 | 0.115 | 0.007 | 0.103 | 0.023 | 1 | |
| 15. % licensed compound (Phase 3) | 0.043 | 0.037 | 0.031 | 0.019 | 0.028 | 0.008 | 0.027 | -0.077 | 0.059 | 0.062 | 0.023 | 0.065 | 0.049 | 0.054 | 1 |

Table 4. Specification Test Results Using CES-Translog production function

| | | Model (1) NONLINEAR OLS | Model (2) GMM |
|--------------|------------------------------------|--------------------------------------|-------------------------|
| Tests | Cobb – Douglas | 1480.7*** | 428.36*** |
| | CES | 1478.1*** | 339.51*** |
| | Translog | 0.45 | 0.18 |
| | Firm Fixed effects | Yes | No |
| | Number of Observations | 748 | 632 |
| | Over-identification test (p-value) | | 0.767 |

- The table reports Chi-Square statistics, with *** denoting p-value < 0.01, related to the Wald test of hypotheses presented in Table 1.

- Both models include the main variables (*R&D* and *Licensing*) and the full set of controls. Model (1) also includes a full set of firm-specific dummy variables. Model (2) estimated with firm fixed-effects did not converge.

- We use *Promotion*, *Trademarks*, their squares, cross product, logs of square terms, and cross-product of logs as instruments for internal and external R&D and the related non-linear terms.

Auxiliary first-stage regressions (OLS linear regressions “within” firm) suggest that the instruments have power. Indeed, the F-test of the joint effect of the instruments on each endogenous variable are 57.97, 5.38, 73.03, 12.57, 7.34 for R&D, the log of R&D-squared, licensing, the log of licensing-squared, and the cross-product of the logs of R&D and licensing, respectively.

Table 5. Panel Regressions. Dependent Variable: log(1+Pipeline).

| | GMM Fixed Effect | | | | Panel RE | Panel FE |
|------------------------------------|--------------------|---------------------|----------------------|---------------------|-----------------------|-----------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Log(R&D) | 1.181 (0.781) | 1.240* (0.671) | 1.078** (0.438) | 0.771** (0.343) | 0.178** (0.0738) | 0.209*** (0.0722) |
| Log(Licensing) | -0.355 (0.306) | -0.143 (0.345) | -0.203 (0.257) | -0.169 (0.180) | 0.0436 (0.0452) | 0.00750 (0.0427) |
| (Log R&D) ² | -0.136 (0.126) | -0.0822 (0.0935) | -0.0696 (0.0709) | 0.0118 (0.0631) | 0.00357 (0.00826) | -0.00945 (0.00868) |
| (Log Licensing) ² | 0.158 (0.127) | 0.244* (0.141) | 0.243** (0.109) | 0.143** (0.0687) | -0.00597 (0.00588) | -0.0107* (0.00607) |
| Log(R&D)* Log(Licensing) | -0.0715 (0.126) | -0.228 (0.162) | -0.221* (0.122) | -0.150 (0.0931) | -0.00110 (0.00935) | 0.0103 (0.00846) |
| Publications | | -0.0853 (0.204) | -0.0428 (0.140) | 0.0251 (0.0923) | 0.0182 (0.0289) | -0.0157 (0.0291) |
| % compound licensed–Phase I | | 1.269 (0.876) | 1.287 (0.786) | 0.823 (0.569) | 0.409* (0.220) | 0.443* (0.229) |
| % compound licensed–Phase II | | -0.137 (0.609) | -0.0646 (0.534) | 0.335 (0.307) | 0.147 (0.152) | 0.212 (0.146) |
| % compound licensed–Phase III | | -0.848 (0.617) | -0.757** (0.378) | -0.523* (0.301) | 0.287* (0.148) | 0.380*** (0.144) |
| Firm Size | | 0.0546* (0.0323) | 0.0487** (0.0244) | 0.0160 (0.0144) | -0.0007 (0.002) | -0.001 (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 | | | Yes | Yes | Yes | Yes |
| Time dummies | | | | Yes | Yes | Yes |
| Observations | 623 | 623 | 623 | 623 | 767 | 767 |
| Log-Likelihood | -556.2 | -693.7 | -676.6 | -470.7 | | -106.3 |
| Cluster | 73 | 73 | 73 | 73 | 92 | 92 |
| Over-identification test (p-value) | 0.530 | 0.753 | 0.736 | 0.116 | | |

- Clustered standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

- While we use the notation “log” we are utilizing the natural log.

- We use the natural logs of *Promotion*, *Trademarks*, *Potential product market size*, *Drug novelty*, *Number of competitors*, their logs of square terms, and cross-product of logs as instruments for internal and external R&D and the related non-linear terms. Auxiliary first-stage regressions (OLS linear regressions “within” firm) suggest that the instruments have power. Indeed, the F-test of the joint effect of the instruments on each endogenous variable are 3.44, 5.05, 1.57, 16.55, 6.85 for R&D, the log of R&D-squared, licensing, the log of licensing-squared, and the cross-product of the logs of R&D and licensing, respectively.

- ATC dummies vary over time because based on the primary ATC defined as the main therapeutic area in terms of annual sales

Table 6. Degree of complementarity/substitution.

| | Panel GMM Fixed Effect | | | | Panel RE | Panel FE |
|-----------------------|------------------------|--------|--------|--------|----------|----------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Mean | -0.021 | -0.031 | -0.029 | -0.019 | -0.0001 | 0.001 |
| Standard Error | 0.033 | 0.033 | 0.024 | 0.017 | 0.001 | 0.001 |
| Median | -0.02 | -0.034 | -0.032 | -0.012 | -0.0001 | 0.0008 |
| Standard Error | 0.033 | 0.041 | 0.031 | 0.011 | 0.001 | 0.0007 |

The table presents estimates of the cross-partial derivative $\frac{d^2 n}{dR_i dR_e}$ (Equation 6) at the mean of the sample.

Table 7a. GMM fixed-effects regressions on split samples based on absorptive capacity levels. Dependent variable: $\log(1+\text{pipeline})$.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|------------------------------------|---------------------------|-------------------------|----------------------|-----------------------|-------------------------|-----------------------|
| | Bottom 25% Publication | Top 25% Publication | Bottom 25% R&D | Top 25% R&D | Bottom 25% Patent | Top 25% Patent |
| Log R&D | 0.299 (0.556) | 1.245*** (0.402) | -1.045 (0.847) | 4.500 (5.962) | 0.448 (0.355) | -0.804 (1.055) |
| Log Licensing | -0.0509 (0.847) | -0.740** (0.341) | -0.0586 (0.395) | -0.260 (1.548) | -1.092** (0.464) | 0.728 (0.735) |
| (Log R&D) ² | -0.0330 (0.0613) | -0.108*** (0.0417) | 0.0978 (0.0956) | -0.353 (0.463) | -0.0205 (0.0553) | 0.0590 (0.130) |
| (Log Licensing) ² | 0.0683 (0.0780) | -0.0394 (0.0518) | -0.142 (0.165) | -0.0761 (0.0761) | 0.348*** (0.0840) | -0.0657 (0.0655) |
| (Log R&D)* (Log Licensing) | -0.00359 (0.107) | 0.155 (0.101) | 0.242 (0.212) | 0.178 (0.286) | -0.272*** (0.0936) | 0.0325 (0.163) |
| Publications | -0.227 (0.159) | -0.488* (0.272) | -0.163 (0.180) | -0.228*** (0.0543) | 0.151 (0.172) | 0.112 (0.305) |
| % compound licensed–Phase I | -0.186** (0.0944) | 0.242 (0.360) | -0.254* (0.143) | -1.658** (0.815) | 0.374*** (0.129) | 0.335 (0.804) |
| % compound licensed–Phase II | 0.325** (0.159) | -2.224*** (0.342) | 0.590*** (0.176) | -1.794*** (0.386) | 1.951*** (0.522) | -1.587*** (0.492) |
| % compound licensed–Phase III | 0.678** (0.328) | -2.203*** (0.456) | 0.713* (0.371) | -2.249*** (0.477) | 0.0621 (0.457) | -0.904 (0.654) |
| Firm Size | -0.0289 (0.228) | -0.0000161 (0.00309) | -0.202 (0.193) | 0.0120 (0.00843) | 0.00452 (0.0305) | -0.00446 (0.00336) |
| Firm Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Time Dummies | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 93 | 124 | 100 | 124 | 104 | 121 |
| Log-Likelihood | -23.80 | 17.64 | -9.087 | 3.005 | -50.59 | 9.551 |
| Cluster | 19 | 19 | 29 | 20 | 25 | 20 |
| Over-identification test (p-value) | 0.692 | 0.330 | 0.186 | 0.700 | 0.244 | 0.444 |

- Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

- While we use the notation “log” we are utilizing the natural log.

Table 7b. GMM fixed-effects regressions on split samples based on economies of scope and licensing experience levels. Dependent variable: $\ln(1+\text{pipeline})$.

| | (1) Bottom 25% Number of ATCs | (2) Top 25% Number of ATCs | (3) Bottom 25% Licensing | (4) Top 25% Licensing |
|---------------------------------------|-------------------------------------|----------------------------------|--------------------------------|-----------------------------|
| Log R&D | 0.309 (0.508) | 0.554** (0.246) | 0.705 (1.440) | 0.113 (0.492) |
| Log Licensing | 0.956 (0.585) | -0.328* (0.176) | 1.311 (2.634) | -1.473 (0.901) |
| (Log R&D) ² | 0.152* (0.0921) | -0.0988** (0.0403) | -0.182 (0.232) | -0.143 (0.0936) |
| (Log Licensing) ² | 0.159 (0.114) | -0.0474 (0.0340) | -0.282 (0.204) | 0.0322 (0.298) |
| (Log R&D)* (Log Licensing) | -0.429*** (0.135) | 0.145** (0.0708) | 0.264 (0.526) | 0.367*** (0.128) |
| Publications | 0.417** (0.187) | -0.216*** (0.0550) | -0.421 (0.437) | -0.0794 (0.195) |
| % compound licensed–Phase I | 2.014*** (0.756) | 0.842 (0.652) | 1.072 (1.251) | -0.671 (0.978) |
| % compound licensed–Phase II | 0.0453 (0.357) | -1.020*** (0.393) | -0.750 (0.888) | 0.0285 (0.291) |
| % compound licensed–Phase III | -0.564*** (0.205) | -0.642 (0.468) | -1.062 (0.928) | -0.415** (0.179) |
| Firm Size | 0.232** (0.0970) | 0.00349 (0.00231) | 0.00940 (0.00858) | 0.0281 (0.203) |
| Firm Fixed Effect | Yes | Yes | Yes | Yes |
| Time Dummies | Yes | Yes | Yes | Yes |
| Observations | 176 | 159 | 111 | 112 |
| Log-Likelihood | -65.55 | 7.785 | -4.450 | 10.30 |
| Cluster | 25 | 20 | 24 | 26 |
| Over-identification test (p-value) | 0.833 | 0.133 | 0.416 | 0.429 |

- Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

- While we use the notation “log” we are utilizing the natural log.

Table 8. Tests on mean complementarity $\left(\frac{d^2n}{dR_i dR_e}\right)$ differences by group of firms.

| | | Full Sample estimation | | Split-sample estimations | |
|----------------------|------------|------------------------|----------------|--------------------------|----------------|
| | | Mean cross-partial | Standard Error | Mean cross-partial | Standard Error |
| Publications | <= 25% | -0.102 | 0.029 | -0.00004 | 0.0002 |
| | > 75% | -0.019 | 0.001 | 0.007 | 0.003 |
| | Difference | 0.082** | | 0.007** | |
| Internal R&D | <= 25% | -0.165 | 0.039 | -0.039 | 0.005 |
| | > 75% | 0.018 | 0.001 | 0.012 | 0.007 |
| | Difference | 0.146*** | | 0.052*** | |
| Patents | <= 25% | -0.063 | 0.026 | -0.032 | 0.026 |
| | > 75% | 0.017 | 0.0006 | 0.059 | 0.004 |
| | Difference | 0.046** | | 0.092*** | |
| Number of ATC | <= 25% | -0.093 | 0.02 | -0.317 | 0.045 |
| | > 75% | -0.045 | 0.008 | 0.0012 | 0.0002 |
| | Difference | 0.047** | | 0.318*** | |
| Licensing experience | <= 25% | -0.218 | 0.039 | -0.077 | 0.016 |
| | > 75% | -0.016 | 0.0005 | 0.019 | 0.002 |
| | Difference | 0.202*** | | 0.09*** | |

- The table contains estimates of the cross-partial $\frac{dn^2}{dR_i dR_e}$ obtained using the GMM estimates from tables 6 and 7.

- The table shows one tail tests.

- ***, **, * indicate that the difference is < 0 at the 0.01, 0.05 and 0.1 confidence levels, respectively.

- 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.

Figure 1. Distribution of complementarity across drivers

