Paper to be presented at the DRUID 2012

on

June 19 to June 21

at

CBS, Copenhagen, Denmark,

Internal R&D Versus In-licensing Complements or Substitutes

Marco Ceccagnoli
Georgia Institute of Technology
College of Management
marco.ceccagnoli@mgt.gatech.edu

Mattew Higgins
Georgia Institute of Technology
College of Management
matthew.higgins@mgt.gatech.edu

Vincenzo Palermo
Georgia Institute of Technology
College of Management
vincenzo.palermo@mgt.gatech.edu

Abstract

This paper analyzes whether internal R&D and external R&D are substitute or complementary strategies with respect to firm innovative performance. Few studies have evaluated this issue, which is surprising since the combination of internal and external R&D investments have been emphasized extensively in the technology strategy literature. We present a two-step model that starts by testing for the proper functional form of the innovation production function. We then use the estimated specification in a profit maximization model to measure the degree of complementarity or substitutability between internal and external R&D. Using a unique dataset, our results suggest that these two types of R&D are complementary activities. Moreover, the degree of complementarity increases over time and it is accentuated for those firms with a higher level of

absorptive capacity, economies of scope and past licensing experience. Our methodological framework can be used by other researchers for a more rigorous understanding of firm and industry characteristics that affect the relationship between internal and external innovative activities.

Jelcodes:L24,-

Internal versus External R&D: Complements or Substitutes?

1. Introduction

Markets for technology have been extensively studied (*e.g.*, Arora and Gambardella 1990 and 1994); however, there still remains a lack of evidence on the determinants of technology demand in terms of the relationship between internal and external R&D (Arora and Gambardella 2009). Firms choose their level of integration but the extent to which they adopt different R&D strategies as substitutes or complements remains uncertain. Some firms, such as Morgan Stanley (2010), have advocated a radical shift for the management of R&D in certain industries. In particular, they argue that the pharmaceutical industry should abandon its current R&D model and fully adopt a "search and development" (S&D) innovation process. Under an S&D framework, firms would abandon all internal research and focus solely on development; 100% of a firm's drug candidates would come from external licensing. While an extreme position, some pharmaceutical companies have openly acknowledged a move towards more external licensing. For example, in 2009 GlaxoSmithKline (GSK) terminated their legendary neuroscience program in order to free up capital to meet their stated goal of allocating 50% of their 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* the marginal or incremental return on the other activity. Complementarity would instead arise if an increase in one of these activities *increased* the marginal returns from the other activity. Substitution between these activities is consistent with the extreme case of total (backward) integration, whereby firms rely exclusively on internal R&D investments and give up the possibility of accessing external know-how. Substitution is also consistent, however, with the opposite case, whereby a fully non-integrated firm relies exclusively on external technology, perhaps yet to be developed, as in the case of the S&D model. The latter model can lead to significant costs related to control and monitoring (Pisano 1990) and IP strategy (Knowles and Higgins 2011). Being completely dependent upon external technology may also alter a firm's relative bargaining position which can also lead to higher overall transaction costs for the firm (*e.g.*, Adegbesan and Higgins 2011).

The choice between these two types of R&D is influenced by whether they are either complements or substitutes which, ultimately, rely on whether synergies exist between them. For example, internal R&D and licensing could fulfill quite distinct yet complementary purposes. R&D may serve functions not directly tied to the creation of new products, such as concept exploration, hypothesis testing, and market credibility. Such activities may be part of routine internal R&D

activities and can complement the investment made on a technology licensed from other firms or institutions.

Indeed, empirical evidence in support of substitution is weak. While Laursen and Salter (2006) find evidence of a substitution effect between these two types of R&D activities, Lowe and Taylor (1998) find complementarity. Complementarity rather than substitution is also supported by the empirical findings of Cassiman and Veugelers (2006) and Tsai and Wang (2008), in the context of internal R&D and external technology acquisitions. However, Vega-Jurado *et al*. (2009) find no evidence of complementarity, nor substitution, in the Spanish manufacturing sector. Their results suggest that firms rely on both internal R&D and external knowledge sources but the two activities do not have synergistic effects.

In light of the mixed evidence on such an important issue for corporate strategy, our goal in this paper is to present an empirical framework that exploits a unique dataset that allows us to provide rigorous firm-level empirical evidence on the substitutability or complementarity relationship between internal and external (specifically, in-licensing) R&D investments. In particular, in order to study the possible synergistic effects between these two types of R&D we adopt a two-step empirical strategy: we first test the coefficients of a flexible CES-Translog innovation production function (Pollak *et al*. 1984) to find the appropriate production function to test complementarity in our context (*e.g.*, Cobb-Douglas, CES, or Translog). We then use these estimates in the context of a profit maximization model to compute the degree of substitutability or complementarity.

Our paper contributes to the literature in several ways. First, by using the proposed framework and focusing on internal R&D and licensing, we find that a complementary relationship exists between them in the context of the global pharmaceutical industry. This is found by estimating the cross-partial effects of these two types of investment on the innovation production function at the firm level, *e.g.*, the effect of increasing one investment on the "physical" marginal returns to the other investment. Second, our approach allows the structural estimation of elasticities of the innovation production function with respect to internal R&D and in-licensing, which represent comparable measures of efficiency of these two types of investments in creating new products. Third, we identify factors conditioning the degree of complementarity. In particular, we find the degree of complementarity increases over time and it is accentuated for those firms with a higher level of absorptive capacity, economies of scope and past licensing experience. Finally, we provide an empirical framework that, conditional on data availability, could be readily applied by researchers within different industry contexts.

We choose the pharmaceutical industry as our research setting because it is an industry where investments in R&D, both internal and external, are a major driver of firm performance (Scherer 2007). Our results, which demonstrate a complementarity between internal and external R&D, provide insight for practitioners suggesting that neither extreme, the purely internal R&D or the S&D model advocated by Morgan Stanley, appears to be the right model for the pharmaceutical industry.

Overall, the implication of our findings suggests that an optimal level of integration between internal and external R&D exists which can have important ramifications on innovative and firm performance.

The remaining paper is organized as follows: Section 2 discusses the relevant literature; Section 3 introduces our theoretical model and empirical methodology. Section 4 and Section 5 discuss the data and our empirical results, respectively. We conclude in Section 6.

2. Literature review

Firms need to 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 industrial organization literature. Early studies recognized and emphasized its implications for the diffusion of technology and product market competition (Shephard 1987; Gallini 1984).

Subsequent literature, for example, has focused on the strength of patent protection and firms' complementary assets as a driver of licensing investments. Gallini (2002) shows that stronger patent protection encourages more efficient knowledge transfer. Arora and Ceccagnoli (2006) find that the effectiveness of patent protection has a stronger impact on the propensity to license when complementary assets are unimportant. These papers complement other factors identified in the literature as affecting the demand side of licensing, including various characteristics of technology, such as its generality (Bresnahan and Gambardella 1998) and economic value (Pisano 1990), the general and abstract nature of underlying knowledge (Arora and Gambardella 1994), and firm size (Arora and Fosfuri 2003; Stuart *et al.* 1999).

Less attention, however, has been paid to the relationship between internal and external R&D (Arora and Gambardella 2010). In most high-tech industries technology buyers conduct extensive internal R&D which may alter their external investment strategy. If so, then this creates a potential tension between developing technology internally and obtaining it externally. This raises the question of whether internal and external R&D are complements or substitutes. A few studies have attempted to address this question but there currently does not appear to be a consensus on whether these two types of R&D are indeed substitutes or complements.

Several complex dynamics can potentially affect this relationship. It is possible that internal R&D activities increase a firm's technological capability to the point that internally developed knowledge is considered superior to external knowledge, giving rise to a "not invented here" syndrome (Allen 1986). In such situations, external ideas, especially those close to a firm's core competency, may be rejected or integrated in the innovative process at a slow rate. When internal efforts are successful this may further reduce the demand for external technology. In this instance, external R&D may be viewed as a substitute for internal R&D.

Pisano's findings (1990) support the substitution viewpoint, but they suggest that this is driven by transaction costs and their influence on the decision to expand R&D externally. Vega-Jurado *et al.* (2009) also provide support for the existence of a substitute relationship. Their research provides empirical evidence on the effect of external knowledge sourcing strategies on the development of both product and process innovation for a sample of innovative Spanish firms. Similarly, Laursen and Salter (2006) find that internal R&D investment negatively moderates the relationship between external knowledge (licensing) and innovation performance. Their findings suggest an inverted U-shaped relationship between the number of external sources a firm utilizes and innovative performance.

Complementarity between internal and external R&D, on the other hand, implies that these two forms of R&D coexist and are interdependent. In contrast to the substitute relationship, this implies that firms which access external R&D must also continue to engage in internal R&D (Chesbrough 2003). Indeed, market-based transactions, such as technology licensing, should produce several benefits that are expected to enhance the upstream innovative activities of the buyer. The first relates to the benefits of specialization at the R&D-level. The buyer, by focusing on its core innovative capabilities and outsourcing the acquisition of competencies in areas in which it is deficient, is expected to improve the marginal productivity of its internal R&D effort. This effect goes beyond the complementarity between the downstream capabilities of the technology buyer, such as development, manufacturing and marketing, and the upstream capabilities of the technology supplier. Such complementarities are perhaps more important and have been the central focus of analysis by scholars in recent years. A consensus seems to have emerged suggesting that they constitute the cornerstone of value creation in the markets for technology (Arora *et al*. 2001).

A second benefit relates to the concept of absorptive capacity (Cohen and Levinthal 1989). Internal R&D investment has a direct effect on the production of new products but it also increases the firm's ability to understand and scan the external technology market, thereby reducing the cost related to the internalization of external know-how. In other words, investment in internal R&D increases the absorptive capacity of the firm (Cohen and Levinthal 1989).

Complementarity has been shown to influence a firm's propensity to access new knowledge (Cassiman and Veugelers 2006; Veugelers and Cassiman 1999; Arora and Gambardella 1994) often through licensing, alliances and acquisitions (Arora and Gambardella 1990, Cockburn and Henderson 1998).

Cassiman and Veugelers (2006) provide empirical evidence on the contextual variables in a firm's strategy which affect the complementary relationship between internal and external R&D. They conclude that the extent to which a firm relies on more "basic" types of know-how affects the strength of the complementarity between innovation activities. Similarly, Tsai and Wang (2008) investigate the extent to which external technology acquisition affects firm performance and how this effect is positively moderated by internal R&D. Their study on Taiwanese electronics manufacturing

demonstrates that external acquisition does not contribute to firm performance *per se* but it has a positive effect when interacting with internal R&D. Finally, in a study on the software industry, Forman *et al.* (2008) find a complementary relationship between internal R&D (*e.g.,* programmers) and external technologies.

While pharmaceutical companies are active in the external technology markets, research focused on the performance of these strategies is mixed. Focusing specifically on licensing, Pisano (1997) argues that a 'lemons' market may actually exist for compounds given the presence of asymmetric information. He finds that licensed compounds are less successful than those developed in-house. While his reasoning is different, Guedj (2009) supports this view. In contrast, Danzon *et al.* (2005) and Arora *et al.* (2009) find the opposite; both studies found that licensed compounds were more successful. Understanding whether internal and external R&D are substitutes or complements may help shed light on these important, unresolved issues.

Other empirical work focusing on downstream performance measures is consistent with a complementarity relationship. The basic idea is that external know-how can quickly bring new resources to the 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 respond to excess capacity generated by patent expirations and to replenish pipeline gaps. Similarly, Chan *et al.* (2007) find that, as a result of downstream complementary assets, firms will engage the external technology market.

Concluding, previous research demonstrates the importance of internal R&D and external technology, especially in our research setting. However, there is still mixed evidence and limited understanding concerning the relationship between these two types of activities as well as its conditioning drivers. In order to clarify these issues, in what follows we propose a framework that can provide structural estimates of the degree of complementarity or substitutability between these two types of R&D investments, as well as their direct effect on a firm's innovation production capability.

3. Model description and estimation procedure

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

We assume each firm is characterized by an innovation production function (n) which depends on the acquisition of external research (R_e), internal R&D expenditure (R_i) and a constant term which represents firm-specific effects as well as other exogenous components of the production function (S):

$$n = f(R_i, R_e, S) \tag{1}$$

Hereafter, the firm and time subscripts are omitted for simplicity. We start by adopting a CES-Translog specification, a flexible functional form which nests the Cobb-Douglas, CES, and Translog specifications (Pollak *et al.*, 1984). As such, it allows us to test which functional form best fits the data.[1] We define our CES-Translog production function as:

$$(2) \qquad \mathrm{Log}(n) = S + \frac{1}{\rho}\ln\left(\alpha_i R_i^{-\rho} + \alpha_e R_e^{-\rho}\right) + \beta_i \ln(R_i)^2 + \beta_e \ln(R_e)^2 + \gamma_{\mathrm{ie}} \ln(\mathrm{R_i}) \ln(\mathrm{R_e}) + u$$

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 even if a firm does not invest in these two type of R&D there is still the possibility to innovate due to the effect of an exogenous component, *S*, which might include factors such as knowledge flows from other firms or universities (Jaffe 1986). 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 βs and γ coefficients represent, respectively, the quadratic impact and the cross-effect of R&D investments on the production of innovations.[2]

The advantage of specification (2) lies in its nested properties. The Cobb-Douglas, CES and Translog forms are all special cases of the CES-Translog. In particular, 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.

<Insert Table 1 here>

3.2. Step 2: Profit maximization under different functional forms

Once we have determined the specification form for the number of innovations produced, we assume that, in equilibrium, firms maximize the expected profit from R&D. However, firms do not know with certainty the actual future value of their investment so we introduce a variable *h*, which

[1] As shown by Equation A.7 in the Appendix, the Cobb-Douglas function does not allow the inputs to be substitutes. We test for the Cobb-Douglas production function in order to have a complete view of the possible specifications available.

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

reflects the firm's expected value from its typical innovation. Following Arora *et al*. (2008), we assume that the expected value of a typical innovation is not influenced by R&D expenditures.

Thus, the final profit function consists of the expected value of an innovation minus total R&D costs and it is described as:

$$\pi = hn - R_i - R_e \tag{3}$$

where *n* is the specification form defined by the results of our tests on the CES-Translog equation. Our empirical findings, discussed below in Section 5.1, will demonstrate that the Translog production function is most appropriate. As such, in the sections that follow, for brevity, we only report the maximization models and the degree of complementarity or substitutability for the Translog production forms. We do, however, provide complete details on the Cobb-Douglas, CES, and CES-Translog models in the Appendix.

3.2.1. Translog function case

The Translog specification extends the Cobb-Douglas case by introducing quadratic terms for the production inputs and is defined as follows:

$$n = \mathrm{R_i^{\alpha_i} R_e^{\alpha_e} e^{S+\beta_i ln[R_i]^2+\gamma_{ie} ln[R_i] ln[R_e]+\beta_e ln[R_e]^2+u}} \tag{4}$$

where *S* is the exogenous component of the production function and R_i and R_e represent internal and external research, respectively.

Substituting Equation (4) into the profit function in Equation (3), we obtain the following first-order conditions:

$$\frac{\mathrm{d\pi}}{\mathrm{dR_i}} = \mathrm{hn}\left(\frac{\alpha_i}{R_i} + \frac{2\beta_i ln[R_i]}{R_i} + \gamma_{ie} ln[R_e]\right) - 1 = \mathrm{hnZ_i} - 1 \tag{5}$$

$$\frac{\mathrm{d\pi}}{\mathrm{dR_e}} = \mathrm{hn}\left(\frac{\alpha_e}{R_e} + \frac{2\beta_e ln[R_e]}{R_e} + \gamma_{ie} ln[R_i]\right) - 1 = \mathrm{hnZ_e} - 1 \tag{6}$$

where Z_i and Z_e represent, for simplicity, the expressions in parentheses. The degree of complementarity or substitutability is thus given by the following second-order condition:

$$\frac{d\pi^2}{dR_i dR_e} = \mathrm{h}\left[\mathrm{n}\left(\frac{\alpha_e}{R_e} + \frac{2\beta_e ln[R_e]}{R_e} + \gamma_{ie} ln[R_i]\right) Z_i + n\gamma_{ie}\right] = \mathrm{hn}(Z_i Z_e + \gamma_{ie}) = \mathrm{hn}\tilde{Z} \tag{7}$$

where $\tilde{Z} = (Z_i Z_e + \gamma_{ie})$ and all other variables are defined above. In contrast to the CES and Cobb-Douglas, the sign of the cross-partial derivative (7) for the Translog functional form is less

intuitive. Although h and n are positive, the sign of $\tilde{Z}$ is ambiguous and we cannot predict ex-ante the sign of the cross-partial derivative. However, we can estimate the predicted value of n and $\tilde{Z}$ for each firm-year by estimating the innovation production function (4). Since we cannot measure h directly, we treat it as a positive unobserved firm specific constant. As a result, although our qualitative estimates of (7) are unaffected, our estimates of the magnitude of the cross-partial derivative (7) are valid only up to an unknown constant that reflects the typical value of a firm's product innovations. This fact creates a concern when comparing the degree of complementarity between groups of firms with different characteristics, as for some of the tests presented in the following sections. As a robustness check, in the empirical analysis we recalculate Equation (7) using the estimated potential market size of a firm's products, available from Pharmaprojects, as a proxy for h. The results presented in this paper remain qualitatively unchanged.

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 substitute the preferred production function into the objective function Equation (3) in order to compute 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 derivatives using Equations (7), (A.4), (A.8) or (A.13).[3]

The model based on the optimization of the objective function, Equation (3), generates 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 potential endogeneity of internal and external (licensing) R&D investments. The source of endogeneity, in this case, 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 potential market size, investments in co-specialized assets and trademarks 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

[3] Equations (A.4), (A.8) and (A.13) are reported in Appendix A.

Our sample is based on a unique longitudinal dataset built from a variety of sources. We started by creating a comprehensive list of global pharmaceutical firms from Pharmaprojects that were active in drug development at any point during the time frame 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 projected market value 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, we obtained product level promotion expenditures from IMS MIDAS™. All financial variables are in 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 with 85% of the firms located in North America and 12% located in Europe and the U.K. The average firm has approximately 11 compounds in their pipeline. Major pharmaceutical companies operate in a number of therapeutic areas. Our firms are no different, operating on average across six therapeutic categories. Almost one third of the compounds under development are focused in three therapeutic areas: Nervous system (ATC=N), Alimentary tract and metabolism (ATC=A) and Cardiovascular (ATC=C).[4]

Dependent variables

Product pipeline. We define our dependent variable as the firm-year product pipeline, which represents a firm's innovative output. The importance of studying a firm pipeline relies on the idea that compounds are developed through different stages and thus need different resources and capabilities in order to reach commercialization. These resources can be developed either internally or acquired through an in-licensing strategy (*e.g.*, Ceccagnoli and Higgins 2011). Using data from Pharmaprojects we generate a yearly pipeline stock by cumulating the number of FDA approved drugs and those in some stage of development (Phase I, II and III) for each firm in our sample.[5] 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), since our objective is to compare the efficiency of internal R&D and in-licensing in obtaining new marketable products.

[4] ATC stands for Anatomical Therapeutic Chemical and it is defined by the World Health Organization (http://www.whocc.no/).

[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 estimation for count data models and our results are confirmed.

Independent variables

Internal R&D investments. We gathered internal R&D data from Compustat and Deloitte Recap which was then used to create a stock variable. R&D data from Compustat includes both internal R&D investment and external acquisitions. In order to isolate just internal R&D we use licensing data from Deloitte Recap and subtract it from the Compustat data. The resulting difference is our proxy for pure internal R&D expenditures. Finally, since developed knowledge can become obsolete or old over time, we use a 15% depreciation rate for this stock variable (Hall 1993).

In-licensing investment (external R&D). We use Deloitte Recap data to collect licensing payments such as milestones, upfront payments and royalties.[6] Our in-licensing variable is based on the sum of milestones and upfront payments. 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]

Instrumental variables

As noted above, internal R&D (R_i) and in-licensing (R_e) may be correlated with unobserved factors. As such, we use potential product market size as an instrument for internal R&D since it captures 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 market size, the more R&D effort is invested 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 drug.

Following Ceccagnoli *et al.* (2009) and Graham and Higgins (2011), we use the value of co-specialized assets as a driver of external R&D expenditures. Previous researchers have found that the possession of downstream complementary assets is a critical driver of innovative performance (Teece 1986; Tripsas 1997; Rothaermel 2001; Chan *et al.* 2007). We use detailing expenditures obtained from IMS MIDAS™ as a proxy for the size of a firm's complementary assets. Larger detailing commitments are suggestive of a more significant complementary asset base.

Finally, our third instrument is the stock of a firm's trademarks. Prior work (Fosfuri *et al.* 2008) has demonstrated their use by firms to reinforce the appropriability of returns from their innovations.

---

[6] We are not able to include royalties in our in-licensing measure because they represent a cost based on future sales, thus, not observable to us. Given this assumption, we acknowledge that our in-licensing measure is downward biased and it represents a lower bound of the in-licensing effect.

[7] Only 9% of data had missing values. We re-estimated the model without the imputed values and results are unchanged.

As such, trademarks could be associated with both internal and external R&D investments. We collect data on active trademarks from the USPTO and build a stock variable to account for intellectual property (IP) accumulation.

Control Variables

We include controls for differences in firm size, technology and product level effects, year effect, average number of therapeutic areas covered, and the percentage of in-licensed compounds in each phase of clinical trial development.

To account for potentially size-driven differences in innovative performance we define *Firm Size* as the total number of firm employees obtained from Compustat (Rothaermel and Boeker 2008). Next, we control for possible differences in uncertainty between in-licensing and internal R&D. Firms that license new compounds may face a higher probability of success since they pay for a compound that has already gone through part of the early stage development process.[8] Using data from Pharmaprojects we define *% of Licensed Compound* as the percent of licensed compounds that a firm has at each phase of the clinical development process.

In order to control for the innovative focus of our focal firm we define *Main Therapeutic Area* as a series of dummy variables covering primary (ATC1) therapeutic categories. 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. In addition, we also control for the total number of therapeutic areas covered, *Number of ATCs*. Finally, we include specifications with time trend dummies and firm-specific dummies 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 and functional form tests

Our estimation procedure starts by estimating the coefficients of a CES-Translog production function (Equation 2). As indicated previously, the tests are summarized in Table 1. The advantage of adopting a flexible and general 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 a 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 we focus

[8] 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. While his reasoning is different, Guedj (2009) supports this view. Arora *et al.* (2009), however, find the opposite. 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.

only on the specification tests described in Table 1. Marginal effects and the degree of complementarity/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 our three different models in Table 4. Model (1) includes only our endogenous variables, Model (2) incorporates all controls, while Model (3), which includes only the endogenous variables, is computed using firm fixed effects.[9]

Our results, which are robust to the entire set of 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 since the tests specified in Table 1 are not 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 and degree of complementarity/substitutability

After identifying Equation (4) as the appropriate production function, we can now estimate marginal effects and the degree of complementarity/substitutability. To facilitate the estimation and the interpretation of coefficients we adopt a log-log form of the Translog production function. This transformation makes the model linear with respect to the natural logarithm of our main independent variables. We then estimate the elasticities and the degree of complementarity/substitutability by taking the derivative with respect to the logarithm of internal R&D and in-licensing (external R&D).

The results of the Translog estimation are reported in Table 5. We use both linear OLS with firm fixed effects and GMM estimation in an attempt to deal with potential endogeneity. Table 6 reports the estimation of the mean cross-partial derivative (Equation 4) as well as the cross-partial derivative at the 25th, 50th, and 75th percentiles (labeled Quartiles 1, 2, and 3, respectively).

<Insert Table 5 and Table 6 here>

Elasticities of innovative output with respect to internal R&D and in-licensing are reported at the bottom of Table 5. In Models (3) and (6), which include all appropriate controls and fixed effects, the elasticities are positive and significant. A percentage increase in internal R&D leads to a percentage increase of 0.124 in Model (3) and 0.372 in Model (6) in the production of innovations, while an

[9] The fixed effect model with all controls did not converge and is thus not reported. The nonlinearity of Equation 4 does not allow us to eliminate the firm fixed effects using first-differences or by transforming the data to within-firm deviations. As a result, we include firm specific dummies and drop the constant term. Details are available from the authors upon request.

increase in licensing expenditure equals a percentage increase of 0.004 and 0.0058 respectively in Model (3) and in Model (6). All these elasticities are significantly different from zero at conventional levels. This finding supports our argument that both factors are important in explaining innovative performance, while suggesting that internal R&D has a much stronger impact on new product introductions.

Overall, these results are consistent with previous studies in the field of industrial organization which show the positive effect of internal and external knowledge sources on innovation (*e.g.*, Ceccagnoli and Higgins 2011; Bresnahan and Gambardella 1998). Moreover, we find evidence of a complementary relationship between internal R&D and in-licensing which is consistent with Cassiman and Veugelers (2006) and Tsai and Wang (2008). To our knowledge, the paper by Tsai and Wang (2008) is the only one that uses in-licensing expenditure to study the relationship between internal and external R&D. While it is directly comparable with our results, they do not use an innovation production function approach.

Results in Table 6 improve upon the existing literature in several ways. First, we use in-licensing investments versus a stock of external deals as a measure of external R&D, which is more common in the extant literature. Second, our empirical results demonstrate a positive effect of both internal R&D and in-licensing on a firm's innovative performance. According to the absorptive capacity framework, internal R&D activities not only increase a firm's technological capabilities but also foster the ability to absorb external knowledge (Cohen and Levinthal 1989). Firms need to integrate internal and external knowledge within their innovation production function in order to capture the synergistic effects each innovative activity has on the return of the other. Logically, higher levels of internal R&D should enhance the ability to integrate externally acquired technologies, thus leading to higher performance. On the other hand, external knowledge compensates for inadequacies in internal capabilities. In-licensing is one way to access new knowledge, which in turn, boosts innovation. The greater the level of internal and external R&D effort, the stronger is their combined effect on a firm's innovative performance.

Internal R&D efforts also improve a firm's exploitation of technological knowledge, its adoption and eventual conversion into innovations (Griliches 1979; Cohen and Levinthal 1989). Investments in internal R&D can increase the effectiveness of external technology acquisition. Empirical evidence suggests that a higher level of R&D effort improves a firm's ability to convert external technical knowledge into innovation activities (Gambardella 1992). R&D investment not only stimulates innovation but also strengthens absorptive capacity to enhance technology acquisition (Griffith *et al*. 2004). Moreover, Song *et al*. (2005) find that internal R&D efforts have a significant impact on the application of technological knowledge.

5.3. Drivers of the complementarity relationship

Our models provide evidence of a complementary relationship between internal R&D and in-licensing. To identify some of the drivers of this relationship we split the sample based on the mean value of four variables: time, stock of internal R&D investment, number of therapeutic categories (ATC) served and stock of in-licensing investments. For each sub-sample we re-estimate the Translog production function (Equation 4) and the degree of complementarity or substitutability, and then we test the difference of the estimated mean cross-partial derivatives for each sub-sample pair. A negative (and significant) value of the resulting difference, presented in Table 7, would imply a higher degree of complementarity for the group above the mean. Our objective is to understand if firms that perform better across the four different drivers experience a different level of complementarity among the two types of investment.

<Insert Table 7 here>

Technology access: According to the markets for technology literature, arms-length transactions for technology have significantly increased over the past few decades. In our setting, this can be traced to the emergence of the biotechnology industry in the late 1980s/early 1990s. Expansion of such markets provides firms with greater access to new technologies and firms have more opportunities to open their boundaries to external knowledge. As pointed out by Arora *et al*. (2001), markets for technology increase the access to external knowledge and to more efficient "customized" technologies. Ideas and new technology can originate both from inside the firm and from the outside market.

Since expanding markets for technology increase the penalty of the "not-invented-here" syndrome (Arora *et al*. 2001), firms are responding by gradually employing more appropriate organizational processes and structures to facilitate technology transactions. In other words, firms are gradually opening up their organizational boundaries, with internal R&D processes being necessary to claim some portion of the value brought by external knowledge (Chesbrough 2003). As such, we expect an increase over time in the degree of complementarity between internal R&D and in-licensing. From this perspective, our results (presented in Table 7) show that, ceteris paribus, the degree of complementarity has indeed significantly increased over time.

Absorptive capacity: Complementarity should also increase for firms with higher absorptive capacity. According to this view, the marginal return on in-licensing increases as the intensity of internal R&D grows, implying that internal know-how will increase the marginal return to external knowledge. This effect can be explained by the absorptive capacity framework introduced by Cohen and Levinthal (1989). Prior knowledge stock is crucial to scan, screen, and absorb external know-how. At the same time, access to external know-how may leverage the efficiency of internal R&D. We measure the level of absorptive capacity by the stock of internal R&D investment. Across both specifications (Table 7) our tests confirm the absorptive capacity hypothesis. Firms with a high level

of internal R&D stocks are characterized, on average, by a stronger complementarity relationship between internal R&D and in-licensing investments.

Economies of scope: Complementarity should increase for firms with a more diversified research pipeline due to economies of scope. We measure pipeline diversity or economies of scope by the number of different therapeutic categories (ATC) in which the firm operates. Scope economies increase productivity when activities can share inputs at little or no additional cost. Knowledge generated in a given technological area may not only be useful in its field, but may also potentially be beneficial to the development of products in other technological areas. If this knowledge can be articulated and codified within the firm (Zollo and Winter 2002) it can then be utilized by other technological areas to improve current development efforts. Our results are confirmed only for the OLS specification, weakly confirming 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 is suggestive that these firms may be using knowledge developed in different fields additively in the innovative process, a view consistent with Henderson and Cockburn (1996).

Prior licensing experience: Complementarity should increase for firms with a larger stock of prior licensing deals. Firms that are highly active in licensing should have refined organizational processes in place to manage and integrate external knowledge (Hoang and Rothaermel 2010). We measure prior experience by the stock of licensing expenditures weighted by their value. In Table 7, across both specifications, we find that higher levels of in-licensing activity enhance the marginal productivity of internal R&D. Moreover, a larger stock of licenses should increase a firm's ability to lower transaction costs with respect to the integration of the new knowledge. As Arora (1996) notes, higher levels of experience in agreement formation facilitates the transfer of tacit knowledge.

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). The extant literature remains unclear about the relationship between these two strategies. Our estimates are consistent with the existence of complementarity between these two types of activities. We suggest that external knowledge acquisition, such as in-licensing, promotes innovation by fostering internal R&D. This view is consistent with other studies which also find evidence of complementarity (Cassiman and Veugelers 2006; Tsai and Wang, 2008). Combined with this other work our results imply that internal R&D is not only associated with greater use of external sources but also suggests a synergistic effect between them. We also analyze possible determinants of this relationship by splitting our sample based on four variables. Our mean tests confirm that complementarity appears to increase over time, and we interpret this as evidence that firms are opening up their boundaries and overcoming the not-invented-here syndrome. Moreover, firms with higher levels of absorptive capacity, alliance experience, and those that are enjoying economies-of-scope are characterized by a higher degree of complementarity.

These results provide support for existing theories and provide insights for further theoretical work on the complementarity of innovative activities. At the same time, we provide a methodological framework that 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 our industry setting and the generalization of our results to other industries, since innovation factors are often determined by industrial dynamics (Malerba 2005). 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 decision of how much to rely on different innovative strategies. Industries which present a different innovative process may experience a different relationship.

A second 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 related to internally developed or externally acquired compounds. For example, Guedj (2009) shows that alliance projects are 21% more likely to move from Phase I to Phase II but co-developed compounds are less successful in later stages (Phase II, Phase III and FDA approval) than internal projects. Conversely, Arora *et al*. (2000 and 2009) suggest that asymmetric information and market imperfections increase costs, and consequentially, the expected value of the licensed compound. 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 its stage of development to account for process development uncertainty.

Finally, while our results help 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 *at 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 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., J. Linn, 2004, "Market size in innovation: Theory and evidence from the pharmaceutical industry," *Quarterly Journal of Economics*, 119(3), 1049-1090.

Allen, T., 1986, "Organizational structure, information technology and R&D productivity," *IEEE Transactions on Engineering Management*, 33, 212-217.

Arora, A., 1996, "Contracting for tacit knowledge: the provision of technical services in technology licensing contracts," *Journal of Development Economics*, 50, 233-256.

Arora, A., M. Ceccagnoli, 2006, "Patent protection, complementary assets, and firms' incentives for technology licensing," *Management Science*, 52(2), 293-308.

Arora, A., M. Ceccagnoli, W. Cohen, 2008, "R&D and the patent premium," *International Journal of Industrial Organization*, 26, 1153-1179.

Arora, A., A. Fosfuri, A. Gambardella, 2001, *Markets for technology: Economics of innovation and corporate strategy*. Cambridge, MA: MIT Press.

Arora, A., A. Fosfuri, 2003, "Licensing the market for technology," *Journal of Economic Behavior and Organization*, 52, 277-295.

Arora, A., A. Gambardella, 1990, "Complementarity and external linkages: the strategies of the large firms in biotechnology," *Journal of Industrial Economics*, 38(4), 361-379.

Arora, A., A. Gambardella, 1994, "The changing technology of technical change: general and abstract knowledge and the division of innovative labor," *Research Policy*, 23, 523-532.

Arora, A., A. Gambardella, 2010, "Ideas for rent: an overview of market for technology," *Industrial and Corporate Change*, 19(3), 775-803.

Arora, A., A. Gambardella, L. Magazzini, F. Pammolli, 2009, "A Breath of Fresh Air? Firm types, scale, scope and selection effects in drug development," *Management Science*, 55(10), 1638-1653.

Arora, A., A. Gambardella, F. Pammolli, M. Riccaboni, 2000, "The nature and the extent of the market for technology in biopharmaceuticals," MPRA Paper No. 15977. http://mpra.ub.uni-muenchen.de/15977/

Bresnahan, T., A. Gambardella, 1998, "Licensing the market for technology," E. Helpman ed. *General Purpose Technologies and Economic Growth*. Cambridge, MA: MIT Press.

Cassiman, B., R. Veugelers, 2006, "In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition," *Management Science*, 52(1), 68-82.

Ceccagnoli, M., S.J.H. Graham, M.J. Higgins, J. Lee, 2010, "Productivity and the Role of complementary assets in Firms' demand for technology innovation," *Industrial and Corporate Change*, 19(3), 839-869.

Ceccagnoli, M., M.J. Higgins, 2011, "Enhancing research productivity through the market for technology," Georgia Institute of Technology, Working Paper.

Chan, T., A.J. Nickerson, H. Owan, 2007, "Strategic management of R&D pipelines with cospecialized investments and technology markets," *Management Science*, 53(4), 667-682.

Chesbrough, H.W., 2003, *Open Innovation: the New Imperative for Creating and Profiting from Technology*. Boston, MA: Harvard Business School Press.

Cockburn, I., R. Henderson, 1998, "Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery," *Journal of Industrial Economics*, 46, 157-182.

Cohen, W.M., D. Levinthal, 1989, "Innovation and learning: The two faces of R&D–Implications for the analysis of R&D investment," *Economic Journal*, 99, 569-596.

Danzon, P., A. Epstein, S. Nicholson, 2004, "Mergers and acquisitions in the pharmaceutical and biotech industries," NBER Working Paper 10536.

Danzon, P.M., S. Nicholson, N. Pereira, 2005, "Productivity in pharmaceutical-biotechnology R&D: the role of experience and alliances," *Journal of Health Economics*, 24, 317-339.

Danzon, P.M., A. Epstein, S. Nicholson, 2007, "Mergers and acquisitions in the pharmaceutical and biotech industries," *Managerial and Decision Economics*, 28(4-5), 307-328.

Forman, C., A. Goldfarb, S. Greenstein, 2008, "Understanding the inputs into innovation: do cities substitute for internal firm resources?" *Journal of Economics and Management Strategy*, 17(2), 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(2), 292-305.

Gallini, N.T., 1984, "Deterrence by market sharing: a strategic incentive for licensing," *American Economic Review*, 74, 931-941.

Gallini, N.T., 2002, "The economics of patents: lessons from recent US patent reform," *Journal of Economic Perspectives*, 16, 131-154.

Gambardella, A., 1992, "Competitive advantages from in-house scientific research: the US pharmaceutical industry in the 1980s," *Research Policy*, 21(5**)**, 391-407.

Gans, J., S. Stern, 2000, "Incumbency and R&D Incentives: Licensing the Gale of Creative Destruction," *Journal of Economics and Management Strategy*, 9(4), 485-511.

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

Graham, S.J.H., M.J. Higgins, 2011, "Pharmaceutical Firm Response to Paragraph IV Challenges: The Role of Complementary Assets and Non-Novel Therapeutics," Georgia Institute of Technology, Working Paper.

Griffith, R., S. Redding, J. Van Reenen, 2004, "Mapping the two faces of R&D: productivity growth in a panel of OECD industries," *Review of Economics and Statistics*, 86(4), 883-895.

Griliches, Z., 1979, "Issues in assessing the contribution of research and development to productivity growth," *Bell Journal of Economics*, 10(1)**,** 92-116.

Guedj, I., 2009, "Ownership Vs. Contract: How Vertical Integration Affects Investment Decisions in Pharmaceutical R&D," SSRN Working Paper 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., I. Cockburn, 1996, "Scale, scope and spillovers: the determinants of research productivity," *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., D. Rodriguez, 2006, "The outsourcing of R&D through acquisition in the pharmaceutical industry," *Journal of Financial Economics*, 80, 351-383.

Hoang, H., F.T. Rothaermel, 2010, "Leveraging internal and external experience: exploration, exploitation, R&D project performance," *Strategic Management Review*, 31(7), 734-758.

Jaffe, A., 1986, "Technological opportunity and spillovers of R&D: evidence from firms' patents, profits, and market value," *American Economic Review*, 76(5), 984-1001.

Knowles, S.M., M.J. Higgins, 2011, "Vertical disintegration in the pharmaceutical industry and the role of IP," *Intellectual Asset Management*, 45*,* 10-15.

Laursen, K., A. Salter, 2006, "Open for innovation: The role of openness in explaining innovative performance among UK manufacturing firms," *Strategic Management Journal*, 27, 131-150.

Lowe, J., P. Taylor, 1998, "R&D and technology purchase through license agreements: complementary strategies and complementary assets," *R&D Management*, 28(4), 263-278.

Malerba, F., 2005, "Sectoral systems. How and why innovation differs across sectors," J. Fagerberg, D. Mowery and R. Nelson, eds. *The Oxford Handbook of Innovation*. Oxford: Oxford University Press.

Morgan Stanley, 2010, *Pharmaceuticals: exit research and create value.*

Nicholson, S., P.M. Danzon, J. McCullough, 2005 "Biotech-pharmaceutical alliances as a signal of asset and firm quality," *Journal of Business*, 78(4), 1433-1464.

Pisano, G. P., 1990, "The R&D boundaries of the firm: an empirical analysis," *Administrative Science Quarterly*, 35, 153-176.

Pisano, G.P., 1997, "R&D performance, collaborative arrangements, and the market-for-know-how: a test of the "lemons" hypothesis in biotechnology," Harvard University, Working Paper.

Pollak, R., R.C. Sickles, T.J. Wales, 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., 2001, "Incumbent advantage through exploiting complementary assets via interfirm cooperation," *Strategic Management Journal*, 22, 687-699.

Rothaermel, F.T., W. Boeker, 2008, "Old technology meets new technology: complementarities, similarities and alliance formation," *Strategic Management Journal*, 29(1), 7-77.

Rothaermel, F. T., M.A. Hitt, L.A. Jobe, 2006, "Balancing vertical integration and strategic outsourcing: Effects on product portfolios, new product success, and firm performance," *Strategic Management Journal*, 27(11), 1033-1056.

Scherer, F.M., 2007, "Pharmaceutical innovation," B. Hall, and N. Rosenberg, eds. *Handbook of the Economics of Technological Innovation*. North Holland.

Shephard, A., 1987, "Licensing to enhance demand for new technologies," *RAND Journal of Economics*, 18, 360-368.

Song, M., H. van der Bij, M. Weggeman, 2005, "Determinants of level of knowledge application: a knowledge based and information-processing perspective," *Journal of Product Innovation Management*, 22(5), 430-444.

Stuart, T.E., H. Hoang, R.C. Hybels, 1999, "Inter-organizational endorsements and the performance of entrepreneurial ventures," *Administrative Science Quarterly*, 44, 315-349.

Teece, D.J., 1986 "Profiting from technological innovation: implications for integration, collaboration, licensing and public policy," *Research Policy*, 15, 285-305.

Tripsas, M., 1997, "Unraveling the process of creative destruction: complementary assets and incumbent survival in the typesetter industry," *Strategic Management Journal*, 18, 119-142.

Tsai, K.H., J.C. Wang, 2008, "External technology acquisition and firm performance: a longitudinal study," *Journal of Business Venturing*, 23, 91-112.

Vega-Jurado, J., A. Gutierrez-Gracia, I. Fernandez-de-Lucio, 2009, "Does external knowledge sourcing matter for innovation? Evidence from the Spanish manufacturing industry," *Industrial and Corporate Change*, 18(4), 637-670.

Veugelers, R., B. Cassiman, 1999, "Make and buy in innovation strategies: evidence from Belgian manufacturing firms," *Research Policy*, 28(1), 63-80.

Zollo, M., S.G. Winter, 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	$\rho = 0$

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
Sales (deflated, Mil. $)	4510.527	10029.110	0	67674.560
Firm size (hundreds)	13.841	26.717	0.001	122
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
1	Product pipeline	1											
2	In-licensing (deflated, Mil. $)	0.492	1										
3	Internal R&D (deflated, Mil. $)	0.628	0.734	1									
4	Detailing stock (deflated, Thousand $)	0.639	0.784	0.937	1								
5	Trademark stock	0.339	0.549	0.668	0.598	1							
6	Expected market size (deflated, Thousand $)	0.131	0.095	0.162	0.115	0.100	1						
7	Sales (deflated, Mil. $)	0.543	0.454	0.738	0.623	0.494	0.254	1					
8	Firm size (hundreds)	0.641	0.495	0.810	0.681	0.543	0.251	0.902	1				
9	Number of ATCs	0.653	0.399	0.501	0.515	0.323	0.206	0.582	0.640	1			
10	% licensed compound (Phase 1)	0.161	0.138	0.131	0.105	0.122	0.167	0.087	0.122	0.103	1		
11	% licensed compound (Phase 2)	0.040	0.070	0.079	0.054	0.037	-0.037	0.081	0.093	0.080	0.013	1	
12	% licensed compound (Phase 3)	0.034	0.033	0.027	0.015	0.024	-0.003	0.053	0.055	0.058	0.044	0.033	1

Table 4. Specification Test Results

		Model (1)	Model (2)	Model (3)
Tests	Cobb – Douglas	0.0001***	0.0001***	0.0001***
	CES	0.0001***	0.0001***	0.0001***
	Translog	0.693	0.676	0.991
	Fixed effects	No	No	Yes

The table presents coefficient tests of Equation 2 and described in Table 1. Model (1) includes only our endogenous variables, Model (2) incorporates all controls, while Model (3), which includes only the endogenous variables, is computed using firm fixed effects.

Table 5. Translog Estimation

	OLS FE estimation			GMM estimation		
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept				-1.958*	-1.560*	-1.048**
				-1.006	-0.809	-0.521
Ln(Internal R&D)	0.507**	0.508**	0.355***	0.758	0.612	0.951***
	-0.203	-0.222	-0.096	-0.543	-0.406	-0.281
Ln(In-licensing)	-0.0779	-0.075	-0.024	-0.375	-0.216	-0.392**
	-0.0709	-0.069	-0.053	-0.451	-0.235	-0.164
Ln(Internal R&D)^2	-0.050**	-0.051*	-0.028**	0.069	0.028	-0.026
	-0.023	-0.026	-0.012	-0.069	-0.051	-0.034
Ln(In-licensing)^2	-0.014**	-0.014**	-0.012*	0.611***	0.251***	0.164**
	-0.006	-0.006	-0.007	-0.122	-0.066	-0.067
Ln(Internal R&D)*Ln(in-licensing)	0.029**	0.029**	0.019*	-0.547***	-0.229***	-0.104
	-0.015	-0.014	-0.011	-0.135	-0.068	-0.066
Firm size		0.004	0.0001		0.001	0.001
		-0.004	-0.002		-0.010	-0.006
Number of ATCs					0.035**	-0.011
					-0.015	-0.015
North America					0.343*	-0.118
					-0.202	-0.246
Europe					0.454*	0.504**
					-0.233	-0.247
% licensed compound (Phase I)		0.385	0.351		-0.830	-0.097
		-0.269	-0.240		-0.559	-0.383
% licensed compound (Phase II)		0.134	0.206		-0.602*	-0.448
		-0.175	-0.169		-0.331	-0.276
% licensed compound (Phase III)		0.282	0.300*		-0.473	-0.718***
		-0.177	-0.167		-0.349	-0.210
RD elasticity	0.083***	0.071***	0.124***	-0.181***	0.207***	0.372***
	-0.006	-0.006	-0.003	*-0.042*	*-0.017*	-0.012
Licensing elasticity	-.007***	-0.013***	0.004**	0.491***	0.121***	0.058**
	-0.002	-0.002	-0.002	*-0.083*	-0.035	-0.024
Main therapeutic areas	No	No	Yes	No	No	Yes
Time dummies	No	No	Yes	No	No	Yes
N	767	767	767	633	633	633

Heteroskedasticity robust standard errors clustered by firm in parenthesis.
Standard errors for elasticities bootstrapped with 1000 reps.
***,**,*: significance at the 0.01, 0.05 and 0.1 confidence levels, respectively.

Table 6. Degree of complementarity/substitution

	OLS FE estimation			GMM estimation		
	(1)	(2)	(3)	(4)	(5)	(6)
Mean	0.118***	0.112***	0.063***	20.730***	3.167***	0.549***
	(0.0046)	(0.0041)	(0.0023)	(1.5052)	(0.1928)	(0.0352)
Quartile 1	0.042***	0.040***	0.025***	-0.242**	0.306***	0.018***
	(0.0015)	(0.0017)	(0.001)	(0.101)	(0.065)	(0.0069)
Quartile 2	0.078***	0.074***	0.044***	4.812***	1.364***	0.175***
	(0.0059)	(0.0056)	(0.0031)	(1.0411)	(0.1028)	(0.0203)
Quartile 3	0.159***	0.149***	0.084***	25.949***	4.129***	0.765***
	(0.008)	(0.0083)	(0.004)	(2.5125)	(0.2837)	(0.076)

The table presents estimates of the cross-partial derivative (Estimation of Equation (7)),
Standard errors bootstrapped with 1000 repetitions

Table 7. Test on the degree of complementarity by splitting the sample

			OLS FE estimation Full model	GMM estimation Full model
Degree of complementarity/ substitutability	Time	<= mean	0.033	0.514
		> mean	0.045	0.826
		t-value	-4.727***	-4.188***
	Internal RD	<= mean	0.034	1.172
		> mean	0.436	4.282
		t-value	-18.62***	-15.52***
	Number of ATC	<= mean	0.002	0.860
		> mean	0.109	0.730
		t-value	-25.715***	1.609
	Licensing experience	<= mean	0.049	0.055
		> mean	0.346	2.065
		t-value	-23.06***	-23.109***

One tail test. ***,**,* the difference is < 0 at the 0.01, 0.05 and 0.1 confidence levels, respectively.

APPENDIX

Cobb-Douglas function case

In the case of the Cobb-Douglas, the production inputs are assumed to be complements by assumption. Innovations are produced according to the following innovation production function:

A.1 $$n = R_i^{\alpha_i} R_e^{\alpha_e} e^{S+v}$$

where *S* stands for the exogenous component of the production function and *v* is random error in innovation. The α coefficients represent the output elasticities on innovation production. More specifically, they measure the responsiveness of the output to a change in levels of the inputs used in production.

These first-order conditions allow us to compute the optimal level of R&D investment:

A.2 $$R_i^* = (h\left(\frac{\alpha_i^2}{\alpha_e}\right)^{\alpha_e} e^{S+v})^{\frac{1}{1-\alpha_i-\alpha_e}}$$

A.3 $$R_e^* = (h\left(\frac{\alpha_e^2}{\alpha_i}\right)^{\alpha_i} e^{S+v})^{\frac{1}{1-\alpha_i-\alpha_e}}$$

Equations (12) and (13) have an interesting implication: both internal R&D and the acquisition of external know-how are positive functions of the exogenous component, *S*, of the production function and the expected value of the typical innovation, *h*.

Under the Cobb-Douglas case, the degree of complementarity or substitutability is given by:

A.4 $$\frac{d\pi^2}{dR_i dR_e} = \mathrm{h}n\frac{\alpha_\mathrm{e}}{\mathrm{R_e}}\frac{\alpha_i}{\mathrm{R_i}} = h\mathrm{nW_iW_e}$$

Equation (14) shows that the sign of the cross-partial derivative is driven by the product between α_i and α_e. Given that the two coefficients are constant and positive, by definition, we prove the assumption that under this specification the two inputs are complements, thus, the Cobb-Douglas specification cannot be used to test for substitutability.

CES function case

While the Cobb-Douglas production function assumes that external R&D is complementary to internal R&D, the CES function allows both investments to be either complements or substitutes. Under the CES specification, the production function is defined as follows:

A.5 $$n = (\alpha_i R_i^{-\rho} + \alpha_e R_e^{-\rho})^{\frac{1}{\rho}} e^{S+v}$$

where α_i and α_e sum to 1, S is the exogenous component of the production function, and R_i and R_e represent internal and external R&D, respectively.

A key parameter is ρ, which represents the elasticity of substitution between R_i and R_e which, by definition, varies between -∞ and +1. As ρ tends to zero the innovation production function approaches the Cobb-Douglas functional form described in the previous section, suggesting a complementarity relationship between the two types of investments.

Given that in equilibrium firms maximize the profit function given by Equation (3) under the CES production function, it follows that the first-order conditions are defined by the following equations:

A.6 $$\frac{d\pi}{dR_i} = hn\frac{\alpha_i}{R_i^{\rho+1}(\alpha_i R_i^{-\rho} + \alpha_e R_e^{-\rho})} - 1 = \mathrm{hnK_i} - 1$$

A.7 $$\frac{d\pi}{dR_e} = hn\frac{\alpha_e}{R_e^{\rho+1}(\alpha_i R_i^{-\rho} + \alpha_e R_e^{-\rho})} - 1 = hnK_e - 1$$

The degree of complementarity or substitutability under the CES specification is given by Equation (18):

A.8 $$\frac{d\pi^2}{dR_i dR_e} = h(\frac{1}{\rho} - 1)\rho\alpha_e\alpha_i R_i^{-\rho-1} R_e^{-\rho-1}(\alpha_i R_i^{-\rho} + \alpha_e R_e^{-\rho})^{\frac{1}{\rho}-2} e^{S+v} = h(1+\rho)\mathrm{nK_iK_e}$$

Recalling that ρ varies between -∞ and +1, the sign of the cross-partial derivative is determined by the following part of Equation (18): $(\frac{1}{\rho}-1)\rho\alpha_e\alpha_i$. We assume that the quantity $\left(\frac{1}{\rho}-1\right)$ is always negative while the sign of $\rho\alpha_e\alpha_i$ remains uncertain.[10]

CES-Translog case

We will adopt the CES-Translog production function if the tests in the first step reject the other specification forms. The production function is defined by Equation (20):

A.9 $n = (\alpha_i R_i^{-\rho} + \alpha_e R_e^{-\rho})^{\frac{1}{\rho}} e^{S+\beta_i \ln[R_i]^2+\gamma_{ie}\ln[R_i]\ln[R_e]+\beta_e \ln[R_e]^2+u}$

To simplify the notation we define

A.10 $G = e^{S+\beta_i \ln[R_i]^2+\gamma_{ie}\ln[R_i]\ln[R_e]+\beta_e \ln[R_e]^2+u}$

This simplification, however, does not affect our computations of the first-order or cross-partial derivatives since *G* is always positive. Substituting Equation (20) in the profit Equation (3) we obtain the following first-order conditions:

A.11 $\frac{d\pi}{dR_i} =$

$$h[G\frac{1}{R_i(\alpha_e R_i^{\rho}+\alpha_i R_e^{\rho})}(\alpha_i R_i^{-\rho}+\alpha_e R_e^{-\rho})^{\frac{1}{\rho}}(-\alpha_i R_e^{\rho}+(\alpha_e R_i^{\rho}+\alpha_i R_e^{\rho})(\gamma_{ie}Ln[R_e]+2\beta_i Ln[R_i]))]-1$$

A.12 $\frac{d\pi}{dR_e} =$

$$h\left[G\frac{1}{R_e(\alpha_e R_i^{\rho}+\alpha_i R_e^{\rho})}(\alpha_i R_i^{-\rho}+\alpha_e R_e^{-\rho})^{\frac{1}{\rho}}\left(-\alpha_e R_i^{\rho}+(\alpha_e R_i^{\rho}+\alpha_i R_e^{\rho})(2\beta_e Ln[R_e]+\gamma_{ie}Ln[R_i])\right)\right]-1$$

[10] It can assume a value equal to 0 only if ρ=1 which implies that the CES function reduces to the following linear function: $n = (\alpha_i R_i + \alpha_e R_e)e^{S+v}$. In the case where ρ=1 the cross-partial derivative is equal to 0.

where *G* is defined by Equation (16). The degree of complementarity or substitutability is then given by:

A.13 $$\frac{d\pi^2}{dR_i dR_e} = h[G\frac{1}{R_i R_e(\alpha_e R_i^{\rho} + \alpha_i R_e^{\rho})^2}(\alpha_i R_i^{-\rho} + \alpha_e R_e^{-\rho})^{\frac{1}{\rho}}(\alpha_e^2 \gamma_{ie} R_i^{2\rho} + \alpha_i^2 \gamma_{ie} R_e^{2\rho} + \alpha_i \alpha_e R_i^{\rho} R_e^{\rho}(1 + 2\gamma_{ie} - \rho) + (\alpha_e R_i^{\rho} + \alpha_i R_e^{\rho})(\alpha_i R_e^{\rho}(2\beta_e Ln[R_e] + \gamma_{ie} Ln[R_i])(-1 + \gamma_{ie} Ln[R_e] + 2\beta_i Ln[R_i]) + \alpha_e R_i^{\rho}(-1 + 2\beta_e Ln[R_e] + \gamma_{ie} Ln[R_i])(\gamma_{ie} Ln[R_e] + 2\beta_i Ln[R_i])))]$$

Similarly to the Translog case, the sign of the cross-partial derivative cannot be determined *a priori*; however, all parameters can be specified, thus it can be estimated.