CESifo WORKING PAPERS

8311
2020

May 2020

Dynamic Effects of Licensing and Knowledge Transfer across Research Stages: Evidence from the Pharmaceutical Industry

Zhili Tian, Ralph Siebert

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8311*

Dynamic Effects of Licensing and Knowledge Transfer across Research Stages: Evidence from the Pharmaceutical Industry

Abstract

In the pharmaceutical industry, firms frequently engage in licensing agreements to overcome innovation challenges and keep up with the pace of developing new drugs. Licensing helps firms jointly develop new drugs and acquire external knowledge, which helps improve their internal drug development capability. Our study examines the dynamic effects of licensing on the success of the licensees' internal drug development across research stages. We adopt a structural equation modeling approach and find that external knowledge transfer via licensing can have differential effects on firms' internal research stage-specific R&D capabilities. The success of transferring external knowledge to an in-licensing firm is critically dependent on the firm's internal R&D capabilities, their financial capability, and the research stage. We find that license agreements formed at the early stage (the discovery and preclinical test phases) and intermediate stage (phase 1 and 2 clinical test phases) exert strong direct and short-run effects on internal R&D capabilities in the same research stages. Moreover, licensing formed at the intermediate stage exerts remarkably strong indirect and long-run effects that impact R&D capabilities in successive research stages. Licensing in the late stage (phase 3 clinical test and product approval) is costly and does little to enhance firms' internal R&D capabilities. Our results also show that licensing is formed among firms with stronger financial resources, and those resources are necessary for a successful technology transfer. Our results also provide insights into the impact of licensing on project success rates across research stages.

JEL-Codes: L240, L250, L650, D220.

Keywords: pharmaceutical drug development, licensing, R&D capabilities, effects of licensing on R&D capabilities.

Zhili Tian
Coastal Carolina University, Department of Management and Decision Sciences
Conway / SC / USA
zhili.a.tian@gmail.com

Ralph Siebert
Purdue University
Department of Economics
Krannert School of Management
West Lafayette / IN / USA
rsiebert@purdue.edu

May 2020

1. Introduction

The pharmaceutical industry faces considerable challenges that impede the innovation process. Studies have shown that only a small fraction of research projects successfully pass the drug development process (see Grabowski and Vernon 1990, Grabowski et al. 2002, DiMasi et al. 1991). The high project failure rate puts enormous pressure on pharmaceutical firms to explore alternative innovation strategies that improve the success of their drug development. R&D licensing agreements are one prominent alternative in which pharmaceutical companies invest. Firms invested nearly $37 billion into licensing (see Cartwright and Ahmed 2017), which makes licensing far more popular than other R&D alliances such as research joint ventures (see EvaluatePharma 2013, Hagedoorn and Narula 1996, Hagedoorn and Duysters 2002).

Licensing is a non-equity R&D partnership that is a core strategy that helps firms jointly develop new drugs and benefit from synergy effects. Pharmaceutical firms use licensing as an instrument to access external knowledge and technologies. More specifically, licensing agreements are formed to treat diseases, define novel drug targets, validate targets, signal transduction pathway know-how, animal models, disease expertise, translational medicine know-how, and biomarkers (Schuhmacher et al. 2016). Companies like GlaxoSmithKline allocated 50% of its R&D budget to external drug development projects (Ceccagnoli et al. 2014). Industry wide, 25% of drugs are developed under licensing terms.

Licensing is an important instrument for firms and can serve multiple purposes: First, licensing provides the opportunity for cooperating firms to pool development resources for launching new research projects completing existing projects. Second, licensing enables firms to acquire external knowledge from their licensing partners. The successful transfer and adoption of external knowledge to boost firms' internal drug development capability is an important aspect that is promoted by many pharmaceutical firms. For example, Pfizer created a new division that concentrated on making externally acquired technologies more accessible to internal scientists (Ceccagnoli et al. 2014).

To date, few scholars have examined whether license agreements help firms transfer external knowledge to improve the licensees' drug development capabilities. Technology adoption and transfer are important, as they determine the licensees' (also referred to as in-licensors') drug development capability and productivity. Our study examines whether licensing agreements help licensees improve their internal drug development capabilities. We concentrate on one type of licensing agreement—ex ante licensing agreements. An ex ante licensing agreement is when firms decide to engage jointly in drug development agreements that are negotiated before the innovation has been explored.

Pisano (1990) has shown that a firm's R&D productivity depends on its internal R&D capability, which describes the ability to develop and exploit new technologies. The internal R&D capabilities determine the success rate of R&D projects, internally developed new products, and R&D intensity (Wang et al. 2008). The primary function of R&D capability is to recombine existing knowledge in novel ways to generate new knowledge or innovations (Cohen and Levinthal 1989, Fleming 2001). Internal R&D capabilities also play a fundamental role in acquiring external technologies and knowledge via licensing. New drug development technologies require firms to establish and invest in internal R&D capabilities such that the externally acquired knowledge can be internally processed and become productive (Pisano 2006). Therefore, firms' internal R&D capabilities form the center of our study when evaluating the effects of licensing on the licensees' internal R&D productivity.

One challenge with evaluating R&D capabilities is that they are not directly observed. Moreover, R&D capabilities in the pharmaceutical industry—which is characterized by a complex market structure—have not been well investigated yet because the drug development process is inherently dynamic. Drugs have to successfully pass multiple research stages and receive the FDA's permission to be marketed. The dynamic drug development process implies that internal R&D capabilities are research stage specific, which has to be considered when evaluating the effects of licensing. Licensing agreements are formed at different research stages and have differential effects depending on when they are formed. This study examines the differential effects of licensing on stage-specific internal R&D capabilities.

The research stage-specific characterization of licensing and R&D capabilities also implies that licensing engagements in one specific research stage can exert: (1) direct and short-run effects on R&D capabilities in that specific research stage that are unmediated by any R&D capability from other stages; and (2) indirect and long-run effects on R&D capabilities in successive stages that are mediated by R&D capabilities in earlier stages. Licensing engagement in one research stage may influence the firm's internal R&D capabilities in the same stage as well as successive stages.

We focus on ex ante licensing where firms engage in R&D partnerships at specific research stages or research phases (discovery, preclinical, phases 1, 2, and 3) of the drug development process. We group firms' engagements in licensing by research phases and examine the licensing effects on research stage-specific internal R&D capabilities. More specifically, we evaluate the licensing effects on firms' internal R&D capabilities using a "structural equation modeling" (SEM) approach. SEM is a useful approach in our context since unobserved R&D capabilities are formulated as latent variables and measured by phase-specific variables. For similar reasons, SEM was adopted in other studies, including Liao et al. (2007), who evaluated the effects of knowledge sharing and absorptive capacity on product innovation capability. Rothaermel and Deeds (2004) estimate an SEM to

evaluate the effects of alliances on firms' product development using the number of approved products and patents as their measured variables.

We estimate a set of linear equations that define the relations between research phase-specific latent variables and observed variables. Our study relates to Halfat (1997) and Kim et al. (2018), who measured R&D capabilities by R&D investment, project portfolio management, and the number of R&D employees. Our model treats the stage-specific R&D capabilities as latent variables, which are measured by variables such as the number of projects in the corresponding research phases, licensing experience, economies of scope (measured by the variety of products in development), and value (scale) variables (Ceccagnoli et al. 2014). Our study specifically addresses the direct and indirect licensing effects across stages. We also allow financial resources to affect internal R&D capabilities and firms' licensing engagements, which are measured by variables such as the number of deals and deal values in the corresponding phases.

Our estimation results confirm that the licensing effects on the licensees' internal R&D capabilities are highly dependent on the R&D phase. Our results show that license agreements formed at the early phases of the drug development process (drug discovery and preclinical test) have a strong direct effect on R&D capability in the discovery phase. That is, the externally acquired R&D activities in the drug discovery and preclinical phases are complementary with the internal R&D activities in the discovery phase. This type of licensing agreement is characterized by pronounced indirect effects that affect the R&D capability in subsequent phases, such as the preclinical stage and phase 1.

Licensing engagements in the intermediate development phases (phase 1 and 2 clinical trials) have significant direct and indirect effects on internal R&D capabilities across all clinical phases. Similar to the early-stage licensing agreements, licensing engagements in phase 1 and 2 clinical trials have strong indirect and long-run effects on R&D capabilities across different phases. Hence, licensing agreements in the early and intermediate stages are characterized by strong indirect effects on R&D capabilities in subsequent phases. These indirect effects are so strong that they frequently dominate the direct effects. The results also show that the indirect effects are more pronounced with firms' financial capabilities.

We also find that licensing agreements formed in the late stage of drug development (phase 3 clinical trials) do not exert any direct or indirect effects on the licensees' late-phase R&D and product approval capabilities. Hence, external and internal R&D capabilities are substitutable in late stages of development.

We also examine the effects of the licensees' improved internal R&D capabilities on their R&D productivity. The results show that licenses strongly improve project success rates (in comparison to non-licensing firms) in early research stages. Licensing formation across research stages is

strongly explained by firms' financial capabilities. Our study shows that the internal and external R&D activities in early and intermediate stages are complements (especially in the drug discovery, preclinical, and phase 1 and 2 stages), but they are substitutes in the late stages of the drug development process (phase 3). However, the success of knowledge transfer critically depends on the strength of the acquirers' internal R&D capabilities, which requires large investments.

Pharmaceutical companies search for alternatives that help improve their internal R&D capabilities and boost their drug development productivity (Paul et al. 2010). The formation of R&D partnerships can help firms pool development funds, share technological know-how, and benefit from synergies. R&D partnerships are a fast and flexible way to get access to complementary resources and skills residing in other companies (Jorde and Teece 2003). In fact, companies are increasingly dependent on R&D partnerships, as the steady increase in the number and scope of R&D partnership suggests (see Feachem and Sachs 2002). DiMasi et al. (1991) reports that higher R&D costs are a big factor underlying the recent trend toward more R&D partnerships.[1] Hagedoorn et al. (2006) mention that R&D partnerships have constituted an important instrument since 1975. During the 1980s, partnerships with small, entrepreneurial biotechnological companies played a crucial role in the development of drugs (see Arora and Gambardella 1990, Barley et al. 1992, Pisano 1990, Powell et al. 1996). In the 1990s, partnerships among large pharmaceutical companies became more frequent. R&D partnerships are quite common, and R&D capabilities have long been of substantial interest to companies, policy makers, and scholars. However, empirical studies on the impact of one specific type of R&D cooperation in the pharmaceutical industry—that is, ex ante licensing agreements—is rather scarce. More insight on this relevant topic is needed. Most existing studies focus on ex post licensing (in which an existing innovation is transferred (see Gallini 1984, Gallini and Winter 1985)), rather than ex ante licensing (in which firms decide to cooperate before the invention has been made).

As we mentioned, licensing become a popular and promising instrument for firms to acquire knowledge and technologies pertaining to new drug development. Firms' incentives to engage in licensing agreements can be determined by the state of their internal R&D capabilities (see Ceccagnoli et al. 2014). A firm's internal R&D capability is its ability to develop and exploit new technologies (see Wang et al. 2008, Cohen and Levinthal 1989, Fleming 2001). A firm's R&D capability can determine the extent to which it can absorb and externally acquire knowledge and technologies. Ceccagnoli et al. (2014) show that licensing becomes more efficient for firms that invest heavily in R&D to generate internal R&D capabilities and absorptive capacity. Jekunen (2014) shows that firms' internal clinical capabilities (the ability to internalize critical external knowledge as well as

[1] The 20 biggest pharmaceutical firms alone formed nearly 1,500 alliances with biotech companies alone from 1997 to 2002.

the means by which the firm organizes developmental activities with key partners) are important success factors for collaborations.

The remainder of the study is organized as follows: Section 2 introduces relevant studies related to the effects of licensing on R&D capabilities. Section 3 introduces the pharmaceutical industry and the licensing effects on research stage-specific R&D capabilities. Section 5 presents the data and defines the variables in our model. Section 6 provides the structural equation model that we use to examine the effects of licensing. Section 7 presents the results of the measurement and structural models. In Section 8, we conclude.

2. Literature Review

Our study is related to the literature on the collaboration of creating innovation in the biotechnology and pharmaceutical industries. In the pharmaceutical industry, lack of knowledge of emerging technologies can have negative impacts on internal R&D capabilities and the development of drugs. In order to acquire external knowledge or R&D capabilities, firms engage in acquisitions, alliances, and licensing (Capron and Mitchell 2009). Licensing plays an important role, as it provides an opportunity for biotechnology firms to gain experience and complementary capabilities related to clinical testing, regulatory filings, and commercialization (Quintana-Garcia and Benavides-Velasco 2004). Our data show that licensing activities in the pharmaceutical industry increased drastically from 1998 to 2011 (see section 5).

The relationship between externally acquired technologies and firms' internal R&D capabilities is a critical success component. Arora and Gambardella (1994) emphasize the relevance of complementarities between externally acquired technologies and firms' internal R&D capabilities. Arora and Gambardella (1994) show that companies need internal know-how to be able to use external know-how effectively. In evaluating the success of cooperations, Cassiman et al. (2005) argue that the relationship between cooperating firms' technologies is important. They show that the merging firms' post-merger R&D activities increase when their technologies are complements, but their R&D investments decrease when their technologies are substitutes. Cassiman and Veugelers (2006) find that internal and external knowledge are complementary. Based on a sample of M&As from the drug, chemical, and electronics industries, Makri et al. (2010) find that complementaries between firms' technological knowledge are relevant contributors that explain post-merger innovation performance. In addition, Makri et al. (2010) find that technology complementarities combined with similarities in the acquirer's and target's knowledge base have positive effects on post-merger innovation. Ceccagnoli et al. (2014) mention that complementarities are better exploited if firms continue investing in internal R&D. Previous research studies (Arora and Gambardella 1994, Ahuja

and Katila 2001, Cassiman et al. 2005, Cassiman and Veugelers 2006, Makri et al. 2010) demonstrate that complementarity between internal and external knowledge is an important determinant for technology transfer.

In a related vein, several studies emphasized the relevance of firms' absorptive capacity; that is, firms' ability to integrate the external knowledge and technologies into their internal R&D. For example, the study by Ceccagnoli et al. (2014) fomulates complementarities as being dependent on absorptive capacity, economies of scope, and licensing experience. If the external R&D activities are complementary to the internal R&D activities, the collaboration can be effective and increase firms' internal R&D capabilities. In contrast, if the external and internal R&D activities are substitutes, the cooperation will act purely as replenishing firms' drug development pipelines, and joint R&D activities will not enhance firms' R&D capability.

Nerkar and Roberts (2004) classify knowledge elements into a firm's proximal knowledge (internal R&D capabilities) and a firm's distal knowledge (external knowledge in our study). Proximal knowledge provides a firm with advantages, making incremental improvements via exploitation, while distal knowledge provides advantages in exploration, which underlies more radical product innovations. Licensing is one opportunity that allows firms to access distal knowledge developed in other firms. Nerkar and Roberts (2004) find that a firm's proximal knowledge increases its new product successes, but they did not find any evidence for the positive relation between a firm's distal knowledge and its new product successes.

In the pharmaceutical industry, Ceccagnoli et al. (2014) show that internal R&D investments and licensing investments are neither complements nor substitutes. Grigoriou and Rothaermel (2017) show that internal investment and external knowledge helped pharmaceutical firms slowly build new biotechnology-related knowledge and adapt to the technological discontinuity. Grigoriou and Rothaermel (2017) tracked the development of internal knowledge of biotechnology from the beginning of the biotechnology revolution (1974) until the end of 1998. They find that external sourcing strategies are less effective when firms can internally generate new knowledge or if they have high internal coordination costs. That is, when the external knowledge and the internally developed knowledge are substitutable, the externally acquired R&D activities become less effective.

In summary, external knowledge and technologies can be complementary or substitutable to the internal knowledge and technologies. The previous literature is inconclusive on the effect of licensing on firms' internal R&D capabilities (some literature finds positive effects, while the other literature finds negative effects).

Our study also addresses the fact that pharmaceutical firms license and acquire external knowledge throughout the various research stages—that is, during the early phases that focus on product development and scientific discovery, the intermediate phases (initial clinical testing phases), and

the late phases that focus on final clinical testing stages (FDA 2018). Licensing of drugs (knowledge or technologies) in one phase may have a different effect on firms' R&D capability than licensing in another phase. Thus, our study examines the differential effect of licensing on firms' R&D capabilities across research stages, as will be detailed in the following sections.

3. The Pharmaceutical Industry

New drug development in the pharmaceutical industry is typically a sequential process. BIO (2016) recorded a total of 9,985 drug observations in clinical drug development and regulatory review phases over a period of 10 years, 2006 to 2015. BIO (2016) assessed the success rate of drug development in four phases: phases 1, 2, 3, and the regulatory filing phase. The overall likelihood of drug approval (LOA) in phase 1 was 9.6%. Table 1 reports the success rates for phases 1-3. Phase 1 has the highest success rate (63.2%), which is explained by the fact that phase 1 testing concentrates on safety aspects, and drug efficacy is not part of the evaluation.[2] The phase 2 success rate (30.7%) is the lowest of the four phases studied, while phase 3 has a success rate of 58.1%. The probability of FDA approval after submitting a New Drug Application (NDA) or Biologics License Application (BLA) was 85.3%. DiMasi et al. (2016) reported very similar success rates, with an overall probability of 11.83% for clinical success. The transition or success probability from phase 2 to phase 3 is 59.52%, from phase 2 to phase 3 is 35.52%, and from phase 3 to regulation filing is 61.95%.

In addition to the low success rates in the clinical phases, these phases are also characterized by high development costs. DiMasi et al. (2003) mentions that costs have been relatively stable in the preclinical phases, but have risen dramatically in clinical phases. Table 1 reports the development cost per phase. The development costs in phase 1 were notably lower than the costs in phases 2 and 3. The ratio of phase 3 to phase 1 development cost is 10.1, the ratio of phase 3 to 2 is 4.4, and the ratio of phase 2 to 1 is 2.3. Clinical development (phases 1-3) accounts for approximately 63% of the costs.

Firms in the pharmaceutical industry reacted to the decrease in R&D productivity and the increase in drug development costs through engaging in mergers and acquisitions, licensing, and other strategic alliances. The benefit of licensing is that pharmaceutical firms spend less compared to acquiring an entire firm.

4. Licensing Engagement and Effects on R&D Capabilities

The pharmaceutical industry experienced challenges in developing new drugs such that pharmaceutical firms faced difficulties in sustaining revenue growth. Additionally, the complexity of drug

[2] Phase 1 success rates may also benefit from delayed reporting bias, as larger companies may not deem failed phase 1 programs as material, and these failures may not be reported publicly.

Table 1 Development Costs and Success Rates

Phases	Development Costs		Success Rates
	mean cost	median cost[1]	(%)
Phase 1	25.3	17.3	63.8
Phase 2	58.6	44.8	30.7
Phase 3	255.4	200.0	58.1

[1] The cost represents average out-of-pocket clinical period costs for investigational compounds (in millions of 2013 U.S. dollars).

discovery and development fundamentally increased. Pharmaceutical firms also experienced pressure to build a new knowledge base during times where R&D expenditures were rising and average sales per drug were falling (Pisano 1996) and (Higgins and Rodriguez 2006). Pharmaceutical firms increasingly invested in internal R&D capability due to the emergence of biotechnology and the associated new drug development technologies; externally acquired technology can be absorbed and properly processed internally (Pisano 2006). In this regard, (Jekunen 2014) show that the success of licensing is dependent on the ability to internalize external knowledge, as well as the means by which the firm organizes developmental activities with key partners.

The drug development process is classified in several research phases (see the FDA), and firms' R&D capabilities are dependent on the development phases (see Paul et al. 2010). In the *discovery stage*, teams of chemists, pharmacologists, and biologists screen thousands of compounds and modify them to fight diseases. In the *preclinical phase*, compounds undergo laboratory and animal testing to answer basic safety questions. The preclinical phase delivers clinical candidate molecules, which have sufficient evidence of biologic activity, sufficient safety, and other drug-like properties (Mohs and Greig 2017).

Clinical research focuses on drug safety and effectiveness. Clinical research is categorized in three phases. In *phase 1* of the clinical trials, drugs are tested on people. Researchers adjust dosing schemes based on preclinical data to find out how much of a drug the body can tolerate and to discover its side effects. In *phase 2*, researchers test a drug on patients to gauge the drug's efficacy and to find side effects. *Phase 3* demonstrates whether or not a drug offers a treatment benefit to a specific population. These trials confirm the efficacy from phase 2 and identify long-term and rare side effects. If a drug is safe and effective, the firm will file an application for marketing the drug (*product approval phase*). The FDA reviews a New Drug Application (NDA) and decides whether to approve or reject the NDA.

Based on the drug development process, we classify firms' internal R&D capabilities into the following six phase-specific R&D capabilities that serve as latent variables in our study: discovery ($CapDiscover$), preclinical ($CapPreClin$), phase 1 ($CapPh1$), phase 2 ($CapPh2$), phase 3 ($CapPh3$), and drug approval ($CapProduct$) capabilities (see Table 2).

Table 2 Latent Variables of R&D Capabilities and Licensing Engagement

Latent Variables	Description
$CapDiscover$ (η_1)	R&D capability in drug discovery phase
$CapPreClin$ (η_2)	R&D capability in preclinical phase
$CapPh1$ (η_3)	R&D capability in phase 1 clinical trial
$CapPh2$ (η_4)	R&D capability in phase 2 clinical trial
$CapPh3$ (η_5)	R&D capability in phase 3 clinical trial
$CapProduct$ (η_6)	Capability in obtaining FDA approval
$EngESLicense$ (η_7)	Engagement in licensing deals in discovery phase and preclinical phase
$EngPh12License$ (η_8)	Engagement in licensing deals in phases 1 and 2 trials
$EngPh3License$ (η_9)	Engagement in licensing deals in phase 3 trials
$CapFin$ (ξ)	Capability in financial support of internal R&D projects and external R&D collaboration projects

Pharmaceutical firms engage in licensing throughout all different research phases, and their effects of licensing are research stage specific. We classify licensing engagements into three groups. We consider licenses formed at (a) the early stages (discovery and preclinical testing), (b) the intermediate stages (phases 1 and 2), and (c) the late stage (phase 3) of the drug development. Our latent variables reflect the licensing engagement at the early stages ($EngESLicense$), the intermediate stage ($EngPh12License$), and the late stage ($EngPhLicense3$).

As mentioned earlier, our study emphasizes the engagement of licensing across research phases and evaluates their direct effects on phase-specific R&D capabilities, as well as their indirect effects on successive research stages through the R&D capabilities in the intermediate phases. Figure 1 displays the effects of licensing and a firm's financial capability on R&D capabilities. It also displays the determinants and effects of licensing engagements. The licensing effects on firms' R&D capabilities are represented by the paths from the latent licensing engagement variables ($EngESLicense$, $EngPh12License$, and $EngPh3License$) to the latent R&D capability variables $CapDiscover$, $CapPreClin$, $CapPh1$, $CapPh2$, $CapPh3$, and $CapProduct$. For example, the effect of licensing in the early stages ($EngESLicense$) on firms' internal R&D capabilities in the discovery phase ($CapDiscover$), the preclinical phase ($CapPreclin$), and phase 1 ($CapPh1$) are depicted by the path coefficients, β_{71}, β_{72}, and β_{73}, respectively.

Moreover, the effect of licensing in the intermediate stages ($EngPh12License$) on firms' internal R&D capabilities in phase 1 ($CapPh1$), phase 2 ($CapPh2$), and phase 3 ($CapPh3$) are illustrated by the path coefficients, β_{83}, β_{84}, and β_{85}, respectively. The path coefficient β_{95} represents the effect of late-stage licensing ($EngPh3License$) on phase 3 capability ($CapPh3$).

Firms need to invest in developing their internal R&D capabilities, which requires a strong financial background. At the same time, licensing engagements require substantial financial resources. Since licensing engagement can crowd out firms' funding resources to invest in promoting internal R&D capabilities, our model specification also includes firms' financial capabilities ($CapFin$) as

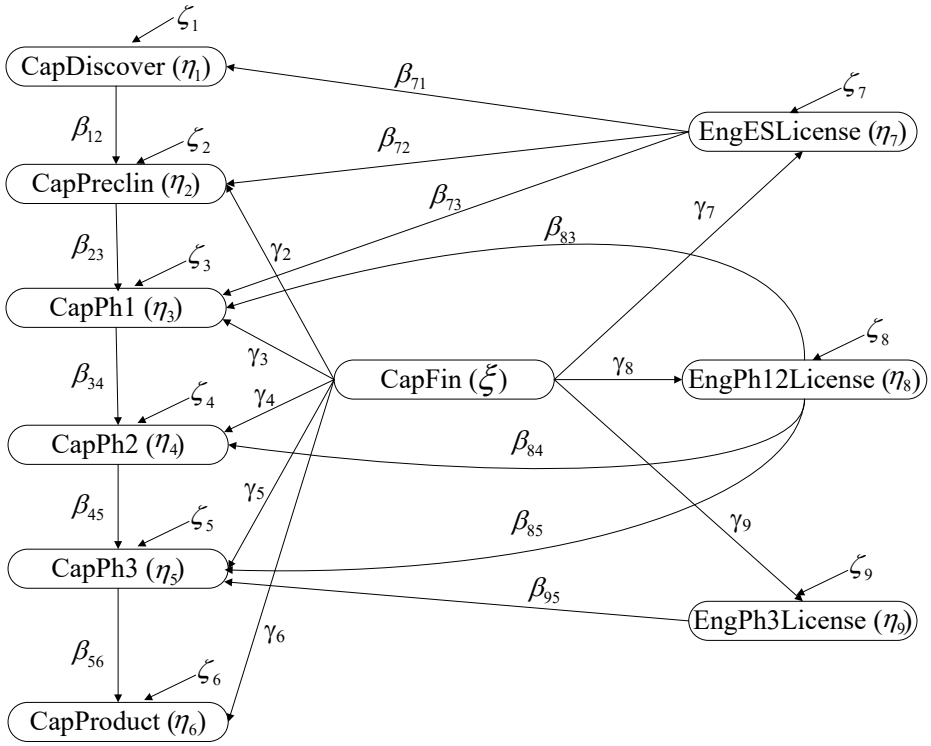


Figure 1 Influence Diagram

a latent variable. As shown in Figure 1, a firm's financial capability ($CapFin$) has an effect on its internal R&D capabilities ($CapDiscover$, $CapPreclin$, $CapPh1$, $CapPh2$, and $CapPh3$, and $CapProduct$), as illustrated by the path coefficients γ_2 - γ_6. Note that firms' financial capability ($CapFin$) does not exert an effect on firms' capability in the discovery phase ($CapDiscover$) since the development cost in this phase is low and this effect is not really explained by firms' revenues and R&D expenditures. Hence, the figure does not show a direct path from $CapFin$ to $CapDiscover$. Firms' financial capability also has an effect on firms' licensing engagements ($EngESLicense$, $EngPh12License$, and $EngPh3License$), as shown by the path coefficients γ_7 - γ_9. Section 7 discusses the licensing effects on firms' R&D capabilities, accounting for direct and indirect effects.

5. Data and Variable Definitions

We use data on licensing and R&D and augment those with financial information for public U.S.-based pharmaceutical and biopharmaceutical firms from 1998 to 2011. The R&D projects and licensing deals are taken from BioPharm Insight, which collects information from U.S. Securities and Exchange Commission filings and a global network of journalists and industry research analysts. The data consist of detailed firm-level licensing deals, research projects, and drug approvals. Our sample includes 311 firms, of which 138 firms engaged in licensing deals.

Firms' licensing engagement is measured by the number of deals and the deal values. Figure 2 displays the number of licensing deals and the total deal values. The number of licensing deals increased from 1998 to 2010, and it reached the peak in 2010, with 275 registered deals. The total deal value also increased in this period, reaching its maximum of $45 billion in 2010.

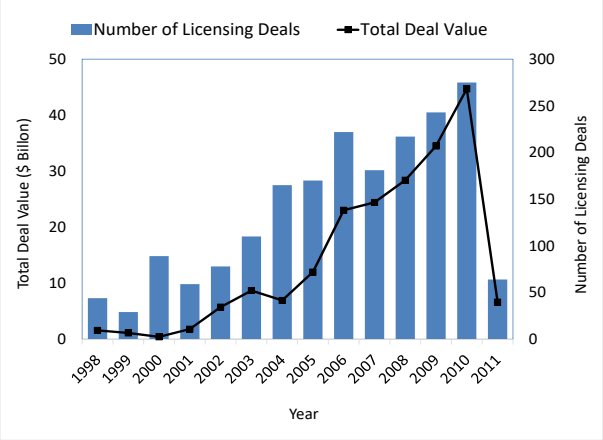


Figure 2 Licensing Deals from 1998 to 2011

Figure 3 depicts the number of licensing deals and the number of NDA and BLA approvals from 1998 to 2011. In 2006 and 2010, we observe sharp reductions in drug approvals, which coincided with a large number of licensing deals. This observation provides some indication that licensing might replenish drug development pipelines; our analysis will devote attention to this possibility.

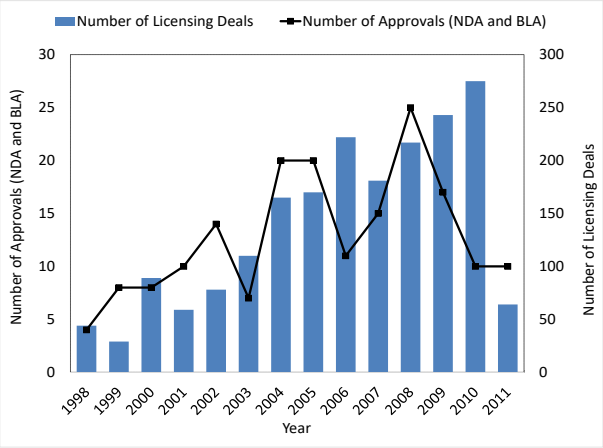


Figure 3 Licensing Deals and New Drug Approvals from 1998 to 2011

Figure 4 shows the R&D spending and licensing deal values. The figure shows that R&D spending increased overall from 1998 to 2011; R&D spending increased from 1998 to 2008and then hovered around $65 billion from 2009 to 2011. Therefore, R&D spending, the number of licensing deals, and total deal values show increasing trends from 1998 to 2010.

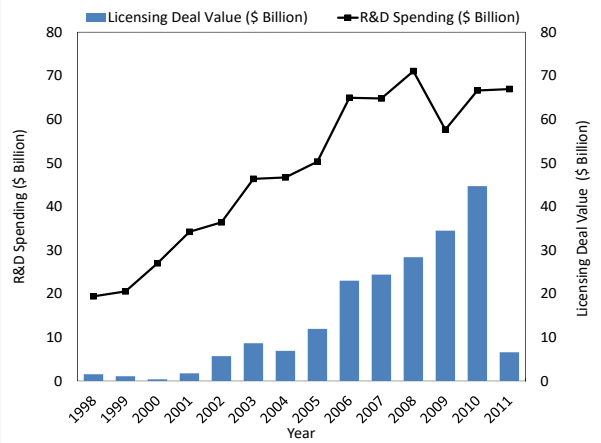


Figure 4 Licensing Deal Value and R&D Spending from 1998 to 2011

Table 3 shows summary statistics for several variables used in our empirical model and classified by licensing and non-licensing firms. The upper panel of Table 3 reports several firm-level characteristics that may determine firms' R&D capabilities and licensing activities. These variables include firms' annual revenues ($Revenues_t$), R&D spending ($RDExpense_t$), the number of licensing deals ($NDeals_t$), and the value of licensing deals ($DealValue_t$). In terms of revenue and R&D investment, licensing firms are characterized by higher annual revenues than non-licensing firms. Licensing firms also invest more in R&D and are involved in more drug development projects than non-licensing firms.

Table 3 Summary Annual Statistics for Licensing and Non-Licensing Firms

	Licensing Firms		Non-Licensing Firms	
	Average	Standard Deviation	Average	Standard Deviation
Revenues ($ billion)	341.2	110.5	11.5	2.4
R&DSpending ($ billion)	52.1	20.7	6.2	2.3
NDeals	145	82		
DealValue ($ billion)	15	15		
NDiscovery	22	30	7	10
NPreclin	31	37	17	21
*NPhase*1	39	42	13	13
*NPhase*2	57	61	18	20
*NPhase*3	30	28	9	10
NApprovals	10	5	3	3

We examine the effect of licensing on licensees' internal R&D capabilities. We build on the study by Ceccagnoli et al. (2014), who formulate complementarities as being dependent on absorptive capacity, economies of scope, and licensing experience. We follow this approach and include measurement variables such as project count variables, project scope variables, licensing deal count and value (scale) variables, and licensing scope variables.

More specifically, we introduce the measurement variables as shown in Table 4. As mentioned above, R&D capabilities in the six research phases ($CapDiscover$, $CapPreClin$, $CapPh1$, $CapPh2$,

$CapPh3$, and $CapProduct$) are treated as latent variables in our study. The corresponding R&D capabilities in year t will be measured by the following measurement variables: 1) The number of projects across research stages: We denote the number of projects in the discovery and preclinical phases, phases 1, 2, and 3, and the product approval stage as $NDiscovery$, $NPreclin$, $NPhase1$, $NPhase2$, $NPhase3$, and $NApprovals$, respectively. 2) The variety of those projects in the corresponding phases: The project variety is measured by the number of projects in each treatment (or indication). Some firms develop drugs in many indications, while others work on only a few indications. We characterize a firm's drug development scope by the number of indications on which a firm works. We consider the number of indications in discovery ($NDiscoverInds$), preclinical ($NPreclinInds$), and phases 1 to 3 ($NPh1Inds$, $NPh2Inds$, and $NPh3Inds$). We denote the number of approved drugs by the FDA as $NApprovals$.

Table 4 Latent Variables and Measurement Variables

Latent Variables	Measurement Variables	Description
$CapDiscover_t$	$NDiscovery_t$	The number of discovery projects in year t
	$NDiscovery_{t-1}$	The number of discovery projects in year $t-1$
	$NDiscoveryInds_{t-2}$	The number of indications projects in year $t-2$
$CapPreClin_t$	$NPreclin_t$	The number of projects in preclinical trial in year t
	$NPreclin_{t-1}$	The number of projects in preclinical phase in year $t-1$
	$NPreclinInds_{t-2}$	The number of indications in preclinical phase in year $t-2$
$CapPh1_t$	$NPhase1_t$	The number of projects in phase 1 in year t
	$NPhase1_{t-1}$	The number of projects in phase 1 in year $t-1$
	$NPh1Inds_{t-2}$	The number of indications tested in phase 1 projects in year $t-2$
$CapPh2_t$	$NPhase2_t$	The number of projects in phase 2 in year t
	$NPhase2_{t-1}$	The number of projects in phase 2 in year $t-1$
	$NPh2Inds_{t-2}$	The number of indications tested in phase 2 projects in year $t-2$
$CapPh3_t$	$NPhase3_t$	The number of projects in phase 3 in year t
	$NPhase3_{t-1}$	The number of projects in phase 3 in year $t-1$
	$NPh3Inds_{t-2}$	The number of indications tested in phase 3 projects in year $t-2$
$CapProduct_t$	$NApprovals_t$	The number of approved drugs in year t
	$NApprovals_{t-1}$	The number of approved drugs in year $t-1$
	$NIndications_{t-2}$	The number of approved drug indications in year $t-2$
$EngESLicense_t$	$LicenseES_t$	The number of licensing deals in discovery or preclinical phases in year t
	$LicESDealVal_t$	The total licensing deal value in discovery or preclinical phases in year t
	$LicESInds_{t-1}$	The number of licensing indications in discovery or preclinical phases in year $t-1$
$EngPh12License_t$	$LicensePh12_t$	The number of licensing deals in phases 1 and 2 in year t
	$LicensePh12_{t-1}$	The number of licensing deals in phases 1 and 2 in year $t-1$
	$LicPh12DealVal_t$	The total licensing deal values in phases 1 and 2 in year t
$EngPh3License_t$	$LicensePh3_t$	The number of licensing deals in phase 3 in year t
	$LicPh3TA_t$	The number of licensing therapeutic areas in phase 3 in year t
	$LicPh3Inds_{t-1}$	The number of licensing indications in phase 3 in year $t-1$
$CapFin_t$	$Revenue_t$	The revenue in year t
	$RDExpense_{t-2}$	The R&D spending in year $t-2$
	$RDExpense_{t-3}$	The R&D spending in year $t-3$

To capture the dynamic effects of licensing and financial capability on the licensees' R&D capabilities, the measurement variables for R&D capabilities in year t include the count variables of

drug development projects and product approvals in years t, $t-1$, and $t-2$. For the same reasons, the measured variables for financial capability and licensing engagement include current and past measurement variables.

Engagement in licensing deals across the research phases ($EngESLicense$, $EngPh12License$ and $EngPh3License$) is formulated as latent variables and proxied by the several measurement variables (see Table 4, lower panel). Engagement in licensing deals in the corresponding research phases will be measured by the number of licensing deals in the corresponding phases ($LicenseES$, $LicensePh12$, and $LicensePh3$). Moreover, since some licensing deals involve large amounts of investments, we also consider the deal values in the development stages ($LicESDealVal$, $LicPh12DealVal$, and $LicPh3DealVal$) The number of licensing deals and deal values measure the scale of the licensing deals.

We also include the number of therapeutic areas of licensing deals in the early stage ($LicESTAs$), phases 1 and 2 ($LicPh12TAs$), and phase 3 ($LicPh3TA$) to characterize the scope of licensing for a firm. Finally, the numbers of indications of licensing deals in the early stage ($LicESInds$), phases 1 and 2 ($LicPh12Inds$), and phase 3 ($LicPh3Inds$) measure the scope of licensing.

Figure 7 in Appendix A.1 summarizes the measurement variables that measure firms' research capabilities, firms' licensing engagements, and firms' financial capability in support of internal R&D and external R&D (licensing) projects.

6. Research Methodology

We evaluate the effect of licensing on firms' R&D capabilities using an SEM approach (see also Liao et al. (2007) and Rothaermel and Deeds (2004)). Based on this technique, we estimate a set of linear equations between latent and measurement variables. In Section 6.1, we introduce the specifications of the measurement model. In Section 6.2, we introduce the specifications of the structural model.

6.1. Measurement Model Specifications

We now introduce the measurement model as illustrated in Fig. 7. The measurement model is based on the measurement variables as shown in Table 4. The specification of the measurement model is illustrated in Figure 7 in the Appendix A.1, where the rectangles represent the measurement variables and the arrows represent the impact of the latent variables on the measurement variables.

Consider the vector of measured variables, $\mathbf{y}$, which is associated with the following latent variables or constructs: $CapDiscover$ (η_1), $CapPreclin$ (η_2), $CapPh1$ (η_3), $CapPh2$ (η_4), $CapPh3$ (η_5), $CapProduct$ (η_6), $EngESLicense$ (η_7), $EngPh12License$ (η_8), and $EngPh3License$ (η_9). The vector $\mathbf{y}$ has 24 elements, $y_1, \cdots, y_{24}$. We let $\mathbf{x}$ be the vector of measured variables associated with the latent variable $CapFin$, which is a 3×1 vector. The measurement model is given by

$$\mathbf{y} = \mathbf{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon}, \tag{1a}$$

$$\mathbf{x} = \mathbf{\Lambda}_x \xi + \boldsymbol{\delta}. \tag{1b}$$

In the above equations, $\boldsymbol{\eta}$ is the vector of latent endogenous variables, which has the following nine elements $\eta_1, \cdots, \eta_9$, ξ is the latent exogenous variable, $\mathbf{\Lambda}_y$ is the coefficient matrix that shows the relation of $\mathbf{y}$ to $\boldsymbol{\eta}$ and has 24 rows and nine columns, $\mathbf{\Lambda}_x$ is the coefficient matrix that shows the relation of $\mathbf{x}$ to ξ and has three rows and one column, and $\boldsymbol{\varepsilon}(24 \times 1)$ and $\boldsymbol{\delta}(3 \times 1)$ are vectors of the measurement errors for $\mathbf{y}$ and $\mathbf{x}$, respectively.

In considering the measurement model in equations (1a) and (1b), the measured variables $\mathbf{y}$ and $\mathbf{x}$ are dependent variables. Thus, unlike in linear regression models, multicollinearity between a count variable in year t and its counterpart in year $t-1$ does not cause parameter estimation problems.

6.2. Structural Model Specifications

We now turn to the structural model to estimate the effects of licensing on a licensee's R&D capabilities, as illustrated in Figure 1. We also consider the fact that the financial capability of a firm (revenue and spending on R&D) determines a firm's R&D capabilities and licensing engagements.

The structural model is defined as follows: Let $\mathbf{B}$ denote the 9×9 coefficient matrix that reflects the effects that the latent endogenous variables have on each other. Some coefficients are zero because we do not include some paths in our model. The nonzero elements of $\mathbf{B}$ are the path coefficients, $\beta_{71}, \cdots, \beta_{91}$ (see Fig. 1). Let $\mathbf{\Gamma}$ be the 9×1 coefficient matrix for the effects of ξ on the vector of endogenous variables, $\boldsymbol{\eta}$. The nonzero elements of $\mathbf{\Gamma}$ are $\gamma_2, \cdots, \gamma_9$ (as shown in Fig. 1). The structural model is given by

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \mathbf{\Gamma}\xi + \boldsymbol{\zeta}, \tag{2}$$

where $\boldsymbol{\zeta}$ is the vector of disturbances, which we assume to have an expected value of zero, $E[\boldsymbol{\zeta}] = \mathbf{0}$, and to be uncorrelated with ξ.

6.3. Direct and Indirect Effects

Using the structural model, we can estimate the direct and indirect effects of licensing and financial capability on R&D capabilities. The direct effect is the effect of one latent variable on another latent variable that is unmediated by any other variables (see Figure 1). The indirect effect of a latent variable is mediated by at least one intervening (intermediate) variable.

To illustrate the direct and indirect effects, we consider the effects of licensing in phases 1 and 2 (latent variable η_8) on a firm's R&D capability in phase 2 (latent variable η_4). The direct effect, represented by β_{84}, is the effect of $EngPh12License$ (η_8) on $CapPh2$ (η_4). The indirect effect works via the intervening variable, the R&D capability in phase 1 (latent variable η_3). A unit change in η_8 results in an expected β_{83} change in η_3. The β_{83} change in η_3 leads to an expected β_{34} change in η_4. The indirect effect of η_8 on η_4 is $\beta_{83}\beta_{34}$. The total effect of η_8 on η_4 is $\beta_{84} + \beta_{83}\beta_{34}$.

7. Model Analyses and Results

We follow Anderson and Gerbing (1988) and estimate the model using a two-step approach; that is, we estimate the measurement model and the structural model. All analyses are based on a covariance matrix of the entire set of measured variables, which we estimate using the maximum likelihood estimation method.

7.1. Examining the Overall Measurement Model

In the measurement model, the sample data are represented by a covariance matrix of measured variables. The measurement model represents the relationships between the latent variables (constructs) and provides an assessment of convergent and discriminant validity. The goodness-of-fit indices for the measurement model measure how well the theoretical covariance matrix matches the covariance matrix of the sample data.

7.1.1. Factor analysis. Confirmatory factor analysis (CFA) tests how well the measured variables represent the latent variables. Measured variables of a specific construct should converge or share a high proportion of variance. High loadings on a factor indicate that the indicators converge to some common point. All factor loadings should be statistically significant. A rule of thumb is that standardized loading estimates should be 0.5 or higher. The factor loadings are shown in Table 7 in Appendix A.2. The table shows that all factor loadings are greater than 0.5 except the ones for $NDiscovery_{t-1}$, $NPreclin_{t-1}$, $NPreclinInds_{t-2}$, $NApprovals_t$, $NApprovals_{t-1}$, and $LicPh12Deals_{t-1}$. A rule of thumb suggests that factor loadings greater than 0.33 are considered to meet the minimal level of practical significance (Ho 2006). By this rule, only $NDiscovery_{t-1}$, $NApprovals_{t-1}$, and $LicPh12Deals_{t-1}$ are considered insignificant (i.e., less than 10% of the variable's total variance is accounted for by their respective factors). However, we keep those measured variables with smaller factor loadings since they appropriately represent the R&D capabilities in drug discovery, product approval, and the engagement in licensing of drugs in clinical tests in phases 1 and 2 ($EngPh12License$). Section 7.1.2 explains why these measured variables should remain in the structural equation model.

7.1.2. Reliability analysis. Reliability is the consistency of measurement of variables for each construct (latent variable). We use Cronbach's alpha reliability measure. This coefficient is an estimate of the average of all the correlation coefficients of the measured variables within a construct. Reliability is also an indicator of convergent validity. The rule of thumb says that an estimate of 0.7 or higher suggests good reliability. A Cronbach's alpha between 0.6 and 0.7 may be acceptable if the other indicators of the model's construct validity are good.

We report the Cronbach's alpha coefficient for each construct and the Cronbach's alpha coefficients after we delete a measurement variable from the construct in Table 7 in Appendix A.2.

If the standardized alpha decreases after removing a variable from the construct, this variable is strongly correlated with other variables in the construct. But, if the standardized alpha increases after removing a variable from the construct, then removing this variable from the construct makes the construct more reliable.

The Cronbach's alpha coefficients are all above 0.63, which demonstrates that all of the measured variables in each construct are reliable, and each construct is internally consistent. The Cronbach's alpha coefficients do not increase, with the exception of removing variables $NPhase2_{t-1}$, $LicPh12Deals_{t-1}$, and $LicPh3Inds_{t-1}$ from their constructs. However, we still keep the three variables in their corresponding constructs for the reasons below. The factor loadings for both—$NPhase2_{t-1}$ and $LicPh12Deals_{t-1}$—are above 0.33, and these variables measure the firm's R&D capability for phase 2 and the engagement in licensing of drugs in phases 1 and 2. Although the factor loading for $LicPh3Inds_{t-1}$ is below 0.33, this variable measures the scope of licensing of drugs in phase 3 for a firm, and it is indispensable for the measurement model.

We now provide some evidence in support of retaining the measured variables with low factor loadings, as identified in section 7.1.1. Removal of $NDiscovery_{t-1}$ will reduce the Cronbach's alpha coefficient for the latent variable $CapDiscover$. Thus, $NDiscovery_{t-1}$ is consistent with other measurement variables that represent the latent variable $CapDiscover$ and should be kept in this construct. Similarly, removal of $NApprovals_{t-1}$ will reduce the Cronbach's alpha coefficient for the latent variable $CapProduct$. Thus, $NApprovals_{t-1}$ is consistent with other measured variables that represent the latent variable $CapProduct$ and should be kept in this construct. Although the removal of $LicPh12Deals_{t-1}$ raises the Cronbach's alpha coefficient for the latent variable $EngPh12License$, we keep this measurement variable because it is an important variable representing the dynamics of licensing activities.

7.1.3. Measurement model estimation. We estimate the measurement model along with the covariance matrix for all latent variables. The measurement model is supported by the various fit indices that are reported in Table 5, column 1. The χ^2 takes on a value of 2244.8 with 359 degrees of freedom. The χ^2 test provides a statistical test of the difference between the covariance matrix of the sample data and the estimated covariance matrix. The smaller the χ^2 statistic, the better the measurement model. However, this statistic increases with the sample size and also increases with the number of measured variables. The ratio of χ^2 value to the degrees of freedom of the model is 6.3. A larger χ^2 may be due to a large sample size (which is 388) and the large number of measured variables (which is 30). Furthermore, in case of large samples, almost every reasonable model will

be rejected if only the chi-square value is considered (Ho 2006). Nevertheless, the measurement model is acceptable based on the ratio of the χ^2 value and the degrees of freedom.[3].

Table 5 Overall Fit Indices

Index	Measurement Model (1)	Structural Model (2)
Chi-Square	2262.6	2270.8
Chi-Square Degree of Freedom	359	382
Standardized RMR (SRMR)	0.16	0.0827
RMSEA	0.117	0.113
RMSEA Lower 90% Confidence Limit	0.108	0.109
RMSEA Upper 90% Confidence Limit	0.121	0.118
Goodness-of-Fit Index (GFI)	0.71	0.70
Bentler Comparative Fit Index	0.82	0.82
Bentler-Bonett NFI	0.79	0.79
Bentler-Bonett Non-normed Index	0.78	0.80

The goodness-of-fit index (GFI) is less sensitive to sample size; it ranges from 0 to 1. Although a higher value of GFI indicates a better fit, no threshold level for acceptability has been established (Ho 2006). Our GFI returns a value of 0.71, indicating a reasonably good model fit given our large sample size.

The root mean square residual is the square root of the squared residuals where the errors are the prediction errors for the elements in the covariance matrix of the sample data. The standardized root mean square residual (SRMS) is 0.158. The root mean square error of approximation (RMSEA) has a known probability distribution. Hence, it represents how well a model fits the population. A lower RMSEA value indicates a better fit. Our RMSEA returns a value of 0.117. The 95% confidence interval of RMSEA is between 0.108 and 0.124. The SRMR and RMSEA along with the 95% confidence interval suggest a good model fit given the large sample size of the data.

The incremental fit indices assess how well a specified model fits relative to a null model, which assumes that all observed variables are uncorrelated. The normed fit index (NFI) is the ratio of the difference in the χ^2 value for the fitted model and a null model divided by the χ^2 value of the null model. A model with perfect fit has an NFI of 1. The Bentler-Bonett NFI and the Bentler-Bonett non-normed index are 0.78 and 0.79, respectively. The Bentler comparative fit index is 0.82. The NFI indices and the comparative fit index all indicate a reasonably good model fit.

We report the factor loadings, variances of error terms, and R squares of the measured variables in Table 8 in Appendix A.2. All the measured variables have statistically significant relationships with their latent variables (constructs) where their t-values exceed the critical value of the 0.1% significance level.

[3] The ratios of the χ^2 values in other studies published in *Management Science*, *Journal of Operations Management*, *Decision Sciences*, and *Journal of Production and Operations Management Society* (in the years from 1984 to 2003) have a mean of 2.17, a median of 1.62, and a range of (0.01, 21.71) (Shah and Goldstein 2006).

7.2. Examining the Overall Structural Model

Section 7.1 demonstrates that the measurement model fits the covariance matrix of the measured variables well. We now evaluate the fit and validity of the structural model.

Table 5 lists the overall fit indices in column 2. Because we restrict some path coefficients to be zero in the structural model, the chi-square of the structural model is higher than the one of the measurement model. The ratio of the χ^2 value to the degrees of freedom of the structural model is 5.9. Thus, the chi-square value indicates that the structural model fits the data well. The GFI is 0.7, which indicates a reasonably good model fit of the structural model given the large sample size.

The SRMR is 0.083, the RMSEA is 0.11, and the 95% confidence interval of the RMSEA is between 0.109 and 0.118. The SRMR < 0.08 and RMSEAs range from 0.05 to 0.08 and are deemed acceptable (Anderson and Gerbing 1988). The SRMS indicates an acceptable model fit. The incremental fit indices are the same as the ones for the measurement model. These indices demonstrate that the structural model fits the covariance matrix of the measured variables well.

The overall structural model replicates the covariance of the measured variables well. We report the factor loadings, variances of error terms, and R-squares of the measured variables in Table 9 in Appendix A.2. All the measurement variables have statistically significant relationships with their latent variables where their t-values exceed the critical value at the 0.1% significance level. The variances of the error terms are all different from zero at the 1% significance level.

The factor loadings for the measured variables in the structural model are roughly the same as those in the measurement model, which are listed in Table 8 in Appendix A.2, even though we reestimate the factor loadings along with the relationship paths in the structural model. Recall that the measurement model assumes that a relationship between each pair of latent variables exists. The factor loading consistency between the measurement model and the structural model demonstrates that the structural model replicates the covariance of the measured variables well, and the structural model represents all relationships among the latent variables.

7.3. Results of the Structural Equation Model

After validating the measurement and structural models in sections 7.1 and 7.2, we now examine the effects of licensing engagements, financial capability, and product approval capability on the licensees' R&D capabilities and licensing activities across various stages, see Figure 5. The direct, indirect, and total effects across all the latent variables are estimated in our structural model. Table 6 reports the estimation results.

Table 6 **The Direct and Indirect Licensing Effects on Product Development and Financial Performance**

From	To		Total	Direct	Indirect
EngESLicense	*CapDiscover*	Effect	0.385	0.374	0.011
		Std Error	0.076	0.073	0.006
		t Value	5.066	5.156	1.910
		p Value	< 0.0001	< 0.0001	0.0561
EngESLicense	*CapPreClin*	Effect	0.357	-0.098	0.455
		Std Error	0.110	0.096	0.100
		t Value	3.238	-1.028	4.568
		p Value	0.0012	0.3039	<.0001
EngESLicense	*CapPh*1	Effect	0.594	0.043	0.551
		Std Error	0.405	0.383	0.169
		t Value	1.467	0.112	3.259
		p Value	0.1425	0.9107	0.00112
*EngPh*12*License*	*CapPh*1	Effect	3.529	3.477	0.052
		Std Error	0.948	0.930	0.031
		t Value	3.723	3.738	1.692
		p Value	0.000197	0.000186	0.0908
*EngPh*12*License*	*CapPh*2	Effect	2.773	-1.024	3.797
		Std Error	0.882	0.475	1.059
		t Value	3.146	-2.154	3.585
		p Value	0.0017	0.0312	0.000337
*EngPh*12*License*	*CapPh*3	Effect	1.081	-0.452	1.533
		Std Error	0.439	0.271	0.504
		t Value	2.461	-1.666	3.043
		p Value	0.0139	0.0958	0.002343
*EngPh*3*License*	*CapPh*3	Effect	-0.210	-0.204	-0.006
		Std Error	0.445	0.432	0.013
		t Value	-0.472	-0.472	-0.462
		p Value	0.6372	0.6372	0.644
*EngPh*3*License*	*CapProduct*	Effect	-0.051	0.000	-0.051
		Std Error	0.109		0.109
		t Value	-0.471		-0.471
		p Value	0.6374		0.6374
CapProduct	*EngESLicense*	Effect	0.348	0.338	0.010
		Std Error	0.198	0.192	0.007
		t Value	1.758	1.763	1.329
		p Value	0.0787	0.078	0.1839
CapFin	*EngESLicense*	Effect	0.050	0.045	0.005
		Std Error	0.003	0.004	0.003
		t Value	16.052	11.360	1.792
		p Value	< 0.0001	< 0.0001	0.0731
CapFin	*EngPh*12*License*	Effect	0.018	0.018	0.000
		Std Error	0.002	0.002	
		t Value	8.350	8.350	
		p Value	< 0.0001	< 0.0001	
CapFin	*EngPh*3*License*	Effect	0.016	0.016	0.000
		Std Error	0.002	0.002	
		t Value	8.529	8.529	
		p Value	< 0.0001	< 0.0001	
CapFin	*CapPreClin*	Effect	0.014	-0.003	0.017
		Std Error	0.004	0.005	0.005
		t Value	3.713	-0.711	3.248
		p Value	0.0002	0.4771	0.0012
CapFin	*CapPh*1	Effect	0.053	-0.034	0.086
		Std Error	0.008	0.025	0.026
		t Value	6.287	-1.328	3.305
		p Value	< 0.0001	0.1841	0.0009
CapFin	*CapPh*2	Effect	0.063	0.025	0.038
		Std Error	0.009	0.009	0.012
		t Value	7.217	2.883	3.336
		p Value	< 0.0001	0.0039	0.0008
CapFin	*CapPh*3	Effect	0.038	0.015	0.024
		Std Error	0.005	0.009	0.009
		t Value	7.822	1.688	2.501
		p Value	< 0.0001	0.0913	0.0124
CapFin	*CapProduct*	Effect	0.014	0.005	0.009
		Std Error	0.002	0.002	0.001
		t Value	7.668	3.245	6.533
		p Value	< 0.0001	0.0012	< 0.0001

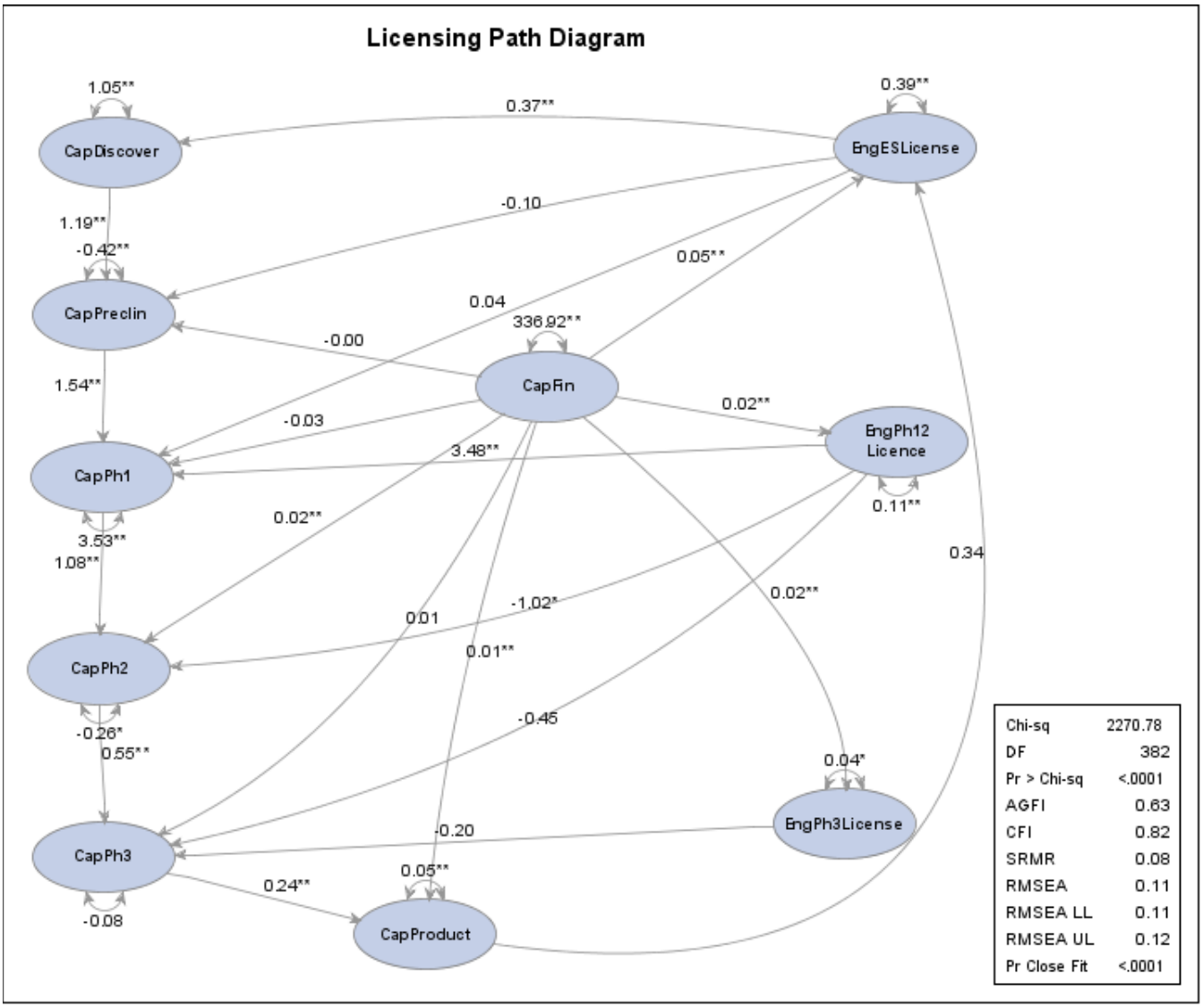


Figure 5 Influence Diagram (Direct Effect)

7.3.1. Effects of licensing on internal R&D capabilities across phases. We begin with presenting the effects of licensing on the licensees' R&D capabilities, as shown in the upper panel of Table 6. Licensing engagements in the discovery and preclinical phases (*EngESLicense*) have positive and statistically significant total effects only on the firms' R&D capabilities in the discovery and preclinical phases. However, the total effect on phase 1 R&D capability is not significant. It is interesting to note that the direct effects of early-stage licensing (*EngESLicense*) are positively significant only for R&D capability in the discovery phase. In contrast, early-stage licensing (*EngESLicense*) indirectly affects the R&D capabilities in the discovery, preclinical, and phase 1 stages. These indirect effects are positive and significant. For example, the indirect effect of early-stage licensing (*EngESLicense*) on R&D capability in the preclinical phase (*CapPreClin*)

becomes effective via the firms' R&D internal capability in the drug discovery phase. Similarly, the indirect effect of early-stage licensing ($EngESLicense$) on phase 1 capability ($CapPh1$) is facilitated by the firms' internal R&D capabilities in the drug discovery and preclinical phases. In general, early-stage licensing ($EngESLicense$) is more frequently characterized by indirect effects than by direct effects on R&D capabilities. This result provides evidence that early-stage licensing has long-run effects that are complementary to in-house R&D capabilities; this type of licensing agreement works via positive indirect effects across different phases. Moreover, it should be recognized that the indirect effect of early-stage licensing ($EngESLicense$) on phase 1 capability ($CapPh1$) is not strong enough to compensate for the insignificant direct effect. This result shows that early-stage licensing improves preclinical research capability, but licensees may not be able to transfer the external preclinical research capability to improve internal R&D capabilities in clinical phases.

Licensing engagements in phases 1 and 2 ($EngPh12License$) have significant effects on phase 1, 2, and 3 capabilities. All direct, indirect, and total effects are significant in this case. Interestingly, the direct effects of licensing on phase 2 and 3 capabilities are negative, and these will be dominated by the positive indirect effects (via the licensee's R&D capabilities in phases 1 and 2), turning the total effects into positive effects. This result shows that even the licensing deals formed in clinical phases are affected by delayed indirect effects that are so strong that they even dominate the negative direct effects. This result also indicates that licensees (often established pharmaceutical firms) acquire technologies from innovators and effectively integrate these external technologies into their R&D. Thus, licensing engagements in clinical trial phases have lagged indirect effects on the licensees' R&D capabilities.

Licensing engagements in the clinical phase 3 ($EngPh3License$) exert no significant total, direct, and indirect effects on firms' late-phase R&D capabilities in phase 3 and the product development phases. Engagement in licensing in the late-stage clinical phases does not influence a firm's R&D capabilities, possibly due to the fact that licensing engagements in the late phase replenish a firm's drug development pipeline.

In summary, licensing agreements at the early research stage have strong indirect effects on R&D capabilities in subsequent phases. This finding highlights the role of firms' absorptive capacities in drug development. That is, firms have to integrate their external technologies into their R&D processes. Licensing at the intermediate stage (phases 1 and 2) has negative direct effects, which are turned into positive total effects due to the strong indirect effects. Licensing at the late stage has no significant direct or indirect effect on licensees' capabilities in the phase 3 clinical tests and the product approval phase.

7.3.2. Effects of product approval and financial capabilities on licensing. We turn to the estimation results of the effects of firms' product approval ($CapProduct$) and firms' spending in R&D or their financial capabilities ($CapFin$) on licensing engagements across phases (see the middle panel of Table 6). A firm's product approval capability ($CapProduct$) has a significantly positive direct effect on early-stage licensing, while the indirect effect is not significant. Firms' financial capabilities ($CapFin$) have significant direct and total effects on licensing across all phases. This finding manifests the important role of financial resources in internal R&D and external collaboration. These results show that financial endowment is a direct predictor for licensing, and it confirms that a strong financial background promotes licensing engagements across all phases. This gives rise to the fact that licensing is associated with high R&D expenses and financially more solid and more capital-intensive firms, such that larger pharmaceutical firms are more inclined to engage in licensing deals. It should be noted that the indirect effects on licensing engagements in clinical phases are close to zero, which further supports the fact that financial strength is necessary for engaging in licensing deals.

7.3.3. Effects of financial capability on pre-clinical and clinical R&D capabilities. The lower panel of Table 6 shows that firms' financial capabilities also exert positive total effects on firms' R&D capabilities across all phases. Having a closer look reveals that the direct effects of a firm's financial capability ($CapFin$) on R&D capabilities are negative but not significant for the preclinical tests and phase 1 clinical tests. This finding shows that even financially more established firms face constraints in drug research and development (discovery and preclinical phases) and, as a solution to this problem, they rely on licensing agreements. The direct effects of a firm's financial capability on R&D capabilities are significant and positive for clinical tests in all phases. The significant and positive direct effects indicate that licensees also invest substantially in developing their own R&D capabilities in clinical phases.

It is remarkable that financial capabilities ($CapFin$) have strong indirect effects on firms' R&D capabilities via licensing activities.[4] The above findings reflect that firms increasingly collaborate with other firms, especially biotechnology firms, to develop drugs in the preclinical and phase 1 to 3 phases.

The positive and significant total effect of financial capabilities on the R&D capabilities in the preclinical and phase 1 test ($CapPreClin$ and $CapPh1$) supports that licensing engagements are

[4] The indirect effect of $CapFin$ on a firm's R&D capability in the preclinical phase works via firms' licensing engagement in the early development stages (discovery and preclinical). Similarly, the indirect effect of $CapFin$ on a firm's R&D capabilities in phases 1 and 2 becomes effective via the firm's licensing in the early stage and intermediate stage (phase 1 and 2 of the clinical tests) of the drug development. The indirect effect of $CapFin$ on a firm's R&D capability in phase 3 works through the firm's licensing in the early stage and intermediate stage of drug development and in phase 3 clinical tests. Similarly, the indirect effect on $CapFin$ on a firm's product approval capability works through licensing in all stages.

helpful in turning the negative direct effects into positive effects. Specifically, financial resources exert significant and positive effects on the above two phases through licensing engagements in previous phases. These firms seem to divert their R&D spending to acquire external knowledge and technologies through licensing in the early stages. This result shows that investment in R&D strengthens firms' research capabilities. The effects of licensing are even more pronounced for the capabilities in the clinical phases 2 and 3 and product approval. Interestingly, the indirect effects of finance on R&D capabilities are significant and positive across all phases. This confirms that R&D capabilities and success in successive research phases depend on research success and capabilities in previous phases. *Moreover, the significant and positive indirect effects also provide evidence that financial resources are crucial for building firms' R&D capabilities (through licensing engagements in the previous phases that affect R&D capabilities in successive phases*). In summary, established firms' financial resources help them develop R&D capabilities in phases 2 and 3 clinical tests and product approvals (i.e., developing expertise in conducting large-scale clinical trials and expertise in working with the FDA to get products approved) both directly and indirectly, but the indirect effects are through licensing engagements.

7.3.4. Project transition rate. After examining the effect of licensing on firms' R&D capabilities, we examine how the enhanced R&D capabilities translate into R&D productivity. The productivity is measured by the transition rate from one phase to a later phase, which we introduce below. After a firm tests a drug in the preclinical phase, it will continue testing the drug in phase 1.

The ratio of the number of projects in phase 1 in year t ($NPhase1_t$) to that of the projects in preclinical phase in year $t-1$ ($NPreclin_{t-1}$) is a proxy of the transition rate from the preclinical phase to phase 1 (S_{01}). That is, $NPhase1_t$ is a function of $NPreclin_{t-1}$,

$$NPhase1_t = S_{01} \times NPreclin_{t-1}.$$

The definitions of transition rates from phase 1 to phase 2 (S_{12}), from phase 2 to phase 3 (S_{23}), and from phase 3 to the FDA approval of drugs (S_{34}) are similar. The following equations relate the number of projects in year t in one phase to the number of projects in year $t-1$ in an earlier phase,

$$NPhase2_t = S_{12} \times NPhase1_{t-1}, \tag{3a}$$

$$NPhase3_t = S_{23} \times NPhase2_{t-1}, \text{ and} \tag{3b}$$

$$NApprovals_t = S_{34} \times NPhase3_{t-1}. \tag{3c}$$

From the structural equations in Section 6, we can derive the transition rates from the above equations. As an example, we derive transition rate S_{12} that equation (3a) defines. From the measurement equation $(1a)$, we have

$$NPhase1_{t-1} = \lambda_8 CapPh1_t + \varepsilon_{8,t-1} = \lambda_8 \eta_{3t} + \varepsilon_{8,t-1}, \tag{4}$$

$$NPhase2_t = \lambda_{10} CapPh2_t + \varepsilon_{10,t} = \lambda_{10} \eta_{8t} + \varepsilon_{10,t}. \tag{5}$$

From the latent variable model (2), we have

$$CapPh2_t = \beta_{34} CapPh1_t + \zeta_{4t} = \beta_{34} \eta_{3t} + \zeta_{4t}. \tag{6}$$

Rearranging terms in equation (4), we have $\eta_{3t} = \frac{1}{\lambda_8} NPhase1_{t-1} - \frac{\varepsilon_{8,t-1}}{\lambda_8}$. Substituting η_{3t} from the above equation into equation (6) yields $CapPh2_t = \beta_{34}(\frac{1}{\lambda_8} NPhase1_{t-1} - \frac{\varepsilon_{8,t-1}}{\lambda_8}) + \zeta_{4t}$. Replacing $CapPh2_t$ in equation (5) yields, $NPhase2_t = \lambda_{10}\left[\beta_{34}(\frac{1}{\lambda_8} NPhase1_{t-1} - \frac{\varepsilon_{8,t-1}}{\lambda_8}) + \zeta_{4t}\right] + \varepsilon_{10,t}$. The coefficient of $NPhase1_{t-1}$ in the above equation is S_{12}. We similarly derive the transition rates S_{01}, S_{23}, and S_{34}. In summary, we estimate the transition rates using the coefficient estimates of the structural equation by the following formulas,

$$S_{01} = \frac{\lambda_7 \beta_{23}}{\lambda_5}, S_{12} = \frac{\lambda_{10} \beta_{34}}{\lambda_8}, S_{23} = \frac{\lambda_{13} \beta_{45}}{\lambda_{11}}, \text{ and } S_{34} = \frac{\lambda_{16} \beta_{56}}{\lambda_{14}}.$$

We calculate the transition rates for firms that engage in licensing using the parameter estimates of the structural equation model in Section 7.2. We also calculate the transition rates for firms that do not engage in licensing using the parameter estimates of the structural equation model that Appendix B provides. For the latter type of firms, we cannot estimate the transition rate from phase 3 ($CapPh3$) to product approval ($CapProduct$) due to the small number of observations associated with this transition.

Licensing firms often possess different R&D capabilities in drug discovery and development than do firms that do not engage in licensing. To eliminate such effect on the transition rates from one phase to the next phase, we normalize the transitions rates S_{01}, S_{12}, S_{23}, and S_{34} by the transition rate from the discovery to the preclinical phase for each firm. Figure 6 displays the normalized transition rates for licensing firms with a solid line and for non-licensing firms with a dashed line.

Firms that engage in licensing have higher transition rates from the early development stage to phase 1 (S_{01}) than the firms that do not engage in licensing (0.78 vs. 0.39). In Section 7.3.1, we find that licensing in early stages of the drug development process ($EngESLicense$) enhances firms' R&D capability ($CapPreClin$). Through licensing, firms (typically large pharmaceutical firms) learn new mechanisms from innovative firms (typically biotechnology firms) for treating diseases or developing drugs. The enhanced R&D capability in the preclinical phase ($CapPreClin$) translates into R&D productivity in the preclinical phase ($CapPreClin$), which results in a higher transition rate of projects from the preclinical tests to phase 1 clinical tests.

Licensing firms have a slightly lower transition rate from phase 1 to phase 2 (S_{12}) than non-licensing firms (0.81 vs. 0.84). The high transition rate S_{12} for both types of firms is also congruent with the industry average transition rate of 59.52% (DiMasi et al. 2003). Although licensing

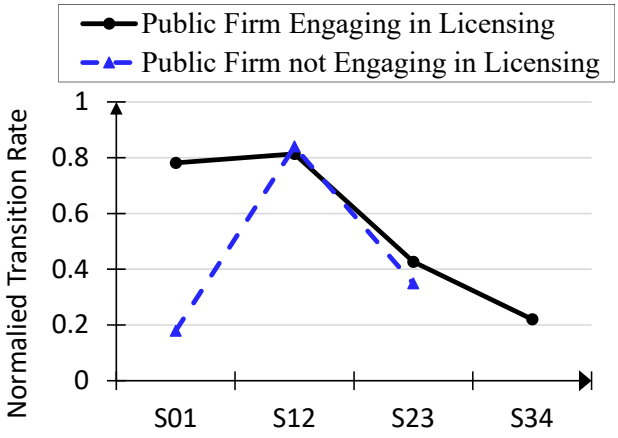


Figure 6 **Project Transition Rates for Firms: Licensing Firms vs. Non-Licensing Firms**

enhances firms' R&D capabilities in phase 1, the enhanced R&D capabilities cannot further improve the project transition rate for licensing firms. This may be because firms already have high transition rates S_{12} and have little room to further improve their productivity.

Transition rates from phase 2 to phase 3 (S_{23}) for both types of firms are lower than those from phase 1 to phase 2. In particular, the transition rates for licensing firms and non-licensing firms are 0.43 and 0.35, respectively. The transition rate for the latter type of firms is the same as the industry average transition rate of 35.52% (DiMasi et al. 2003). Phase 3 clinical trials are expensive, and firms have difficulties in finding patients. Facing financial constraints in R&D, some firms discontinue their projects in phase 3. However, licensing in phases 1 and 2 enhances firms' R&D capabilities in phase 2 ($CapPh2$). In addition, licensing in the early stage of drug development ($CapESLicense$) enhances firms' R&D capabilities in phase 1 ($CapPh1$). The knowledge gained and technologies acquired from licensing in the early stage and phase 1 carry over to phase 2. The enhanced R&D capabilities translate into higher productivity for licensing firms (a 7% higher transition rate) than for the non-licensing firms.

Because project transition rates from phase 2 to phase 3 for both types of firms are low, the transition rate S_{34} from phase 3 to approved products is also low. In fact, only licensing firms have a sizable transition rate from phase 3 to product approval. Licensing in phases 1 and 2 indirectly enhances firms' R&D capability in phase 3. The enhanced R&D capability in turn boosts the productivity for licensing firms from other firms. Thus, licensing firms have a sizable transition rate S_{34}. Although licensing in phases 1 and 2 indirectly enhances firms' R&D capabilities in phase 3, licensing in phase 3 does not enhance R&D capability in phase 3. Since licensing in phase 3 does not increase firms' R&D capabilities, the transition rate from phase 3 to product approval S_{34} for licensing firms is lower than the industry average of 61.95% (DiMasi et al. 2003).

8. Discussion and Conclusions

Pharmaceutical firms face enormous pressure to develop new drugs. In order to keep up with this innovation pressure, firms often engage in R&D alliances such as licenses, which become increasingly important for the development of new drugs. Licensing allows firms to acquire external knowledge and expertise from other firms that is beneficial for the development of new drugs. This knowledge and expertise are dependent on the specific drug development phases. The successful transfer of externally acquired knowledge to (inside) the firm is critically dependent on firms' internal R&D capabilities, which are specific to the drug development phases.

Our study examines the effects of licensing on firms' internal R&D capabilities across drug development phases and on firms' product approval capabilities. The evaluation is challenging because the R&D capabilities are not observable. To overcome this challenge, we adopt a structural equation modeling approach that enables us to examine the direct, indirect, and total effects of licensing on firms' internal R&D capabilities and product approval capability. We also examine the direct, indirect, and total effects of firms' financial capabilities on their R&D capabilities, licensing engagement, and product approval capability.

Our estimation results show that licensing can be a useful instrument for firms to access partner firms' knowledge and technologies. However, the successful transfer of externally acquired knowledge is critically dependent on how well the acquirer's internal R&D capabilities are developed. It also depends on the development stage when two firms form their licensing agreements. Licensing formed in the early and intermediate stages of the drug development process improve licensees' internal R&D capabilities in the discovery, preclinical, phase 1, and phase 2 stages. Our results provide evidence that licensing can have indirect and long-lasting effects.

More specifically, licensing in the early stage of the drug development process enhances the licensee's R&D capability in phase 1, and licensing engagement in the intermediate stage of drug development (phases 1 and 2) raises internal R&D capabilities in phases 2 and 3. The indirect effects of early-stage licensing on the acquirer's internal R&D capabilities are large, and they support the fact that absorptive capacity plays an important role in transferring external knowledge in-house. Our results also show that financial capabilities facilitate licensing and improve internal R&D capabilities.

Licensing in later stages is costly and does not do much to enhance firms' internal R&D capabilities. Our study stresses that the internal and external R&D activities in early stages are complements (especially in the drug discovery, preclinical, and phase 1 and 2 stages), but they are substitutes in the late stage (phase 3) of the drug development process.

To conclude, our study suggests that licensing can be an appropriate instrument to help learning and acquiring external knowledge. Especially at the early stages, externally acquired knowledge

via licensing can be complementary to internal R&D capabilities. However, the success of knowledge transfer critically depends on the strength of the acquirers' internal R&D capabilities, which requires large investments. Therefore, the success of licensing and the transfer of external knowledge to boost internal drug development is critically dependent on the state of firms' internal R&D capabilities and their financial strength.

References

Ahuja, G., R. Katila. 2001. Technological acquisitions and the innovation performance of acquiring firms: a longitudinal study. *Strat. Mgmt. J.* **22** 197–220.

Anderson, J.C., D.W. Gerbing. 1988. Structural equation modeling in practice: a review and recommended two-step approach. *Psychological Bulletin* **103**(3) 453–460.

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

Arora, A., A. Gambardella. 1994. Evaluating technolological information and utilizing it: Scientific knowledge, technological capability, and external linkages in biotechnology. *Journal of Economic Behavior and Organization* **24** 91–114.

Barley, S., J. Freeman, R. Hybels. 1992. Strategic alliances in commercial biotechnology. N. Nohria, R. Eccles, eds., *Networks and organizations: Structure, form, and action*. Harvard Business School Press, Boston,MA, 311–347.

BIO. 2016. Clinical development success rates 2006-2015. Biotechnology Innovation organization (BIO), Washington, DC, BioMedTracker, CA, and Ampion, OR. `http://go.bio.org/rs/490-EHZ-999/images/Clinical%20Development%20Success%20Rates%202006-2015%20-%20BIO%2C%20Biomedtracker%2C%20Amplion%202016.pdf`.

Capron, L., W. Mitchell. 2009. Selection capability: How capability gaps and internal social frictions affect internal and external strategic renewal. *Organization Science* **20**(2) 294–312.

Cartwright, H., T. Ahmed. 2017. Ims pharmadeals: Review of 2016. Tech. rep., QuintilesIMS.

Cassiman, B., M.G. Colombo, P. Garrone, R. Veugelers. 2005. The impact of m&a on the r&d process: An empirical analysis of the role of technological and market relatedness. *Research Policy* **34** 195–220.

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., M. Higgins, V. Palermo. 2014. Behind the scenes: Sources of complementarity in r&d. *Journal of Economics & Management Strategy* **23**(1) 125–148.

Cohen, W.M., D.A. Levinthal. 1989. Innovation and learning: The two faces of r&d. *The Economic Journal* **99**(397) 569–596.

DiMasi, J.A., H.G. Grabowski, R.W. Hansen. 2016. Innovation in the pharmaceutical industry: New estimates of r&d costs. *Journal of Health Economics* **47** 20–33.

DiMasi, J.A., H.G. Grabowski, R.W. Hansen, L. Lasagna. 1991. Cost of innovation in the pharmaceutical industry. *Journal of Health Economics* **10** 107–142.

DiMasi, J.A., R. Hansen, H.G. Grabowski. 2003. The price of innovation: New estimates of drug development costs. *Journal of Health Economics* **22** 151–185.

EvaluatePharma. 2013. World preview 2013, outlook to 2018. Tech. rep., Evaluate Ltd.

FDA. 2018. The drug development process. `https://www.fda.gov/patients/learn-about-drug-and-device-approvals/drug-development-process`.

Feachem, R.G.A., J.D. Sachs. 2002. Global public goods for health. Tech. rep., Report of Working Group 2 of the Commission on Macroeconomics and Health, WorldHealth Organization, Geneva.

Fleming, L. 2001. Recombinant uncertainty in technological search. *Management Sci.* **47**(1) 117–132.

Gallini, N.T. 1984. Deterrence by market sharing: A strategic incentive for licensing. *The American Economic Review* **74**(5) 931–941.

Gallini, N.T., R.A. Winter. 1985. Licensing in the theory of innovation. *The RAND Journal of Economics* **16**(2) 237–252.

Grabowski, H., J. Vernon, , J.A. Dimasi. 2002. Returns on research and development for 1990s new drug introductions. *PharmacoEconomics* **20** 11–29.

Grabowski, H., J. Vernon. 1990. A new look at the returns and risks to pharmaceutical r&d. *Management Science* **36** 804–821.

Grigoriou, K., F.T. Rothaermel. 2017. Organizing for knowledge generation: internal knowledge networks and the contingent effect of external knowledge sourcing. *Strat. Mgmt. J.* **38**(2) 395–414.

Hagedoorn, J., G. Duysters. 2002. Learning in dynamic inter-firm networks: The efficacy of multiple contacts. *Organization Studies* **23**(4) 525–548.

Hagedoorn, J., R. Narula. 1996. Choosing organizational modes of strategic technology partnering: International and sectoral differences. *Journal of International Business Studies* **27**(2) 265–284.

Hagedoorn, J., N. Roijakkers, H. Kranenburg. 2006. Inter-firm r&d networks: The importance of strategic network capabilities for high-tech partnership formation. *British Journal of Management* **17** 39–53.

Halfat, C.E. 1997. Knowhow and asset complementarity and dynamic capability accumulation: the case of r&d. *18* **5** 339–360.

Higgins, M.J., D. Rodriguez. 2006. The outsourcing of r&d through acquisitions in the pharmaceutical industry. *Journal of Financial Economics* **80**(2) 351–383.

Ho, R. 2006. *Handbook of Univariate and Multivariate Data Analysis and Interpretation with SP*. Chapman and Hall/CRC.

Jekunen, A. 2014. Decision-making in product portfolios of pharmaceutical research and development managing streams of innovation in highly regulated markets. *Drug Design, Development and Therapy* **2014**(8) 2010–2016.

Jorde, T.M., D.J. Teece. 2003. Competition and cooperation: Striking the right balance. *California Management Review* **31** 25–37.

Kim, M., J. Kim, Y. Sawng, K. Lim. 2018. Impacts of innovation type smes r&d capability on patent and new product development. *Asia Pacific Journal of Innovation and Entrepreneurshipl* **12**(1) 45–61.

Liao, S., W. Fei, C. Chen. 2007. Knowledge sharing, absorptive capacity, and innovation capability: an empirical study of taiwans knowledge- intensive industries. *Journal of Information Science* **33**(3) 340–359.

Makri, M., M.A. Hitt, P.J. Lane. 2010. Complementary technologies, knowledge relatedness, and invention outcomes in high technology mergers and acquisitions. *Strat. Mgmt. J.* **31**(6) 602–628.

Mohs, R.C., N.H. Greig. 2017. Drug discovery and development: Role of basic biological research. *Alzheimer's & dementia (New York, N. Y.)* **3**(4) 651–657.

Nerkar, A., P. Roberts. 2004. Technological and product-market experience and the success of new product introductions in the pharmaceutical industry. *Strat. Mgmt. J.* **25**(8/9) 779–799.

Paul, S.M., D.S. Mytelka, C.T. Dunwiddie, C.C. Persinger, B.H. Munos, S.R. Lindborg, A.L. Schacht. 2010. How to improve r&d productivity: the pharmaceutical industry's grand challenge. *Nature* **9** 203–214.

Pisano, G. 1990. The r&d boundaries of the firm: an empirical analysis. *Administrative Science Quarterly* **35**(1) 153–176.

Pisano, G. P. 1996. *The Development Factory: Unlocking the Potential of Process Innovation*. Harvard Business School Press, Boston,MA.

Pisano, G. P. 2006. *Science Business: Promise, Reality, and the Future of Biotechnology.*. Harvard Business School Press, Boston,MA.

Powell, Walter W., Kenneth W. Koput, Laurel Smith-Doerr. 1996. Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly* **41**(1) 116–145.

Quintana-Garcia, C., C. A. Benavides-Velasco. 2004. Cooperation, competition, and innovative capability: a panel data of european dedicated biotechnology firms. *Technovation* **24** 927–938.

Rothaermel, F.T., D.L. Deeds. 2004. Exploration and exploitation alliances in biotechnology: a system of new product development. *Strat. Mgmt. J.* **25**(3) 201–221.

Schuhmacher, A., O. Gassmann, M. Hinder. 2016. Changing r&d models in research-based pharmaceutical companies. *Journal of Translational Medicine* **14** 1–11.

Shah, R., S.M. Goldstein. 2006. Use of structural equation modeling in operations management research: Looking back and forward. *Journal of Operations Management* **24** 148–169.

Wang, C., I. Lu, C. Chen. 2008. Evaluating firm technological innovation capability under uncertainty. *Technovation* **28** 349–363.

Appendix A: Measurement Model Specification, Estimates of Measurement and Structural Models, and Transition Rates

This appendix provides the specification of the measurement model and estimates of the measurement and structural models for public firms with in-licensing. The appendix also provides the derivation of the transition rates from one phase to the next phase.

A.1. Measurement Model Specification

This appendix provides the specification of the measurement model as illustrated in Fig. 7. The rectangles represent the measurement variables. The arrows in the figures represent the impact of the latent variables on the measurement variables.

A.2. Estimates of the Measurement and Structural Models

Table 7 in this appendix reports the factor loadings in the CFA and the Cronbach's alpha coefficients in the reliability analysis. Tables 8 and 9 in the appendix report the factor loadings, the variance of error terms, and the R-squares of the measured variables in the measurement and structural models.

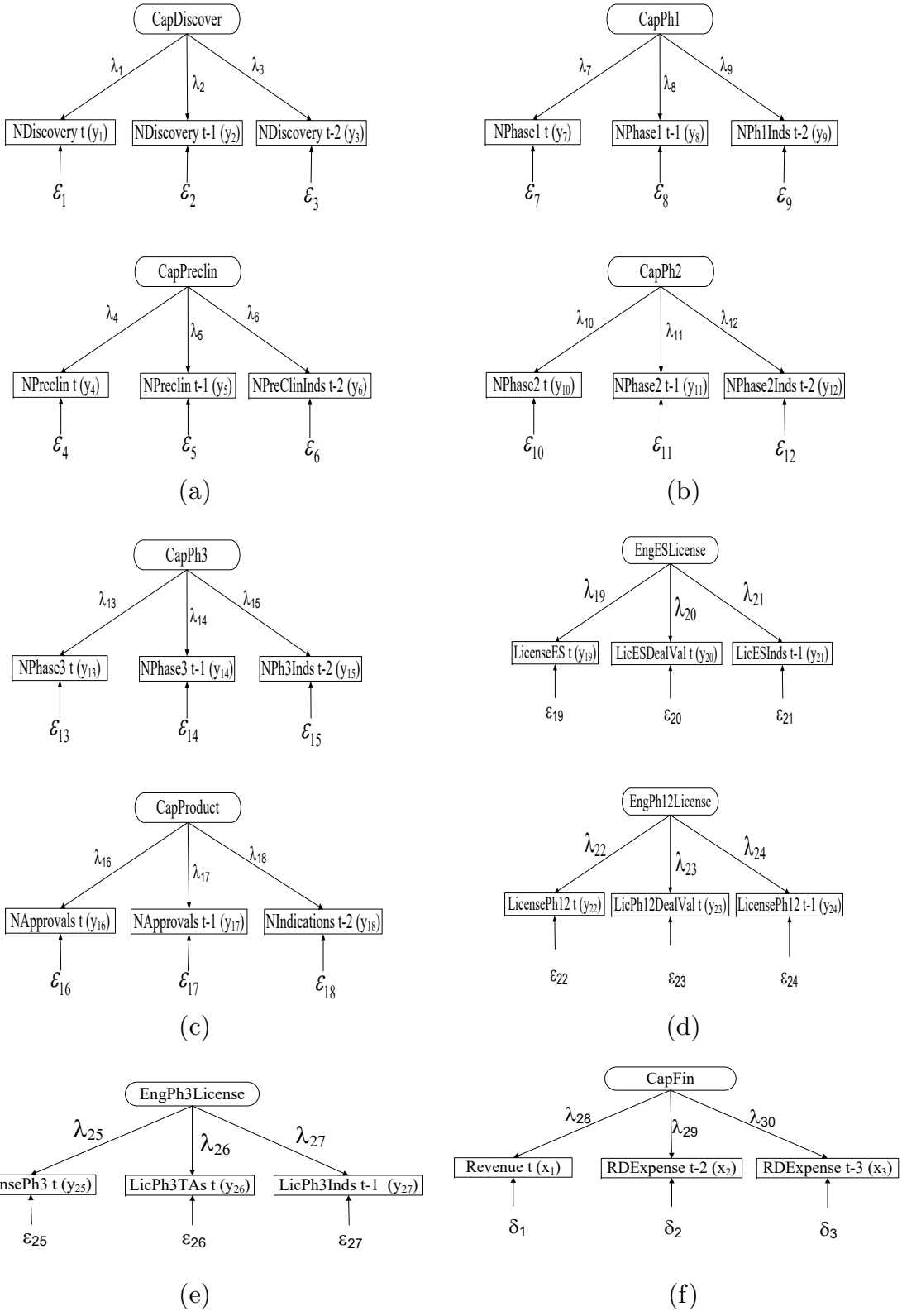


Figure 7 Measurement Models: R&D Capabilities, Licensing Engagement, and Financial Capability

Table 7 Factor Analyses and Cronbachs Alpha Coefficients

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Commun-ality	Cronbachs Alpha Coefficient	
								with All Variable	without One Variable
$NDiscovery_t$	0.2564	0.1521	0.1699	0.0086	**0.7243**	0.0511	0.6963	0.7278	0.5945
$NDiscovery_{t-1}$	0.4312	0.0933	0.1032	-0.0477	**0.1943**	-0.0120	0.7224		0.7066
$NDiscovery_{t-2}$	0.2546	0.0415	0.5002	-0.0109	**0.4783**	0.0382	0.6249		0.6154
$NPreclin_t$	0.1499	0.0488	0.0020	-0.0780	**0.6724**	0.0348	0.4921	0.6395	0.6322
$NPreclin_{t-1}$	0.2655	0.0419	0.2432	0.0052	**0.3332**	-0.0107	0.6797		0.5144
$NPreclinInds_{t-2}$	0.5671	0.0504	0.0824	-0.0789	**0.3212**	0.0739	0.6739		0.4689
$NPhase1_t$	**0.7130**	0.1286	0.2830	0.0386	0.2925	0.3370	0.8252	0.9063	0.8826
$NPhase1_{t-1}$	**0.7157**	0.0836	0.4638	-0.0046	0.1493	0.1896	0.8854		0.8721
$NPh1Inds_{t-2}$	**0.8029**	0.0917	0.3177	-0.0148	0.1888	0.0431	0.8083		0.8419
$NPhase2_t$	**0.8361**	0.1952	0.1199	-0.0635	0.1359	0.1036	0.8047	0.8689	0.8028
$NPhase2_{t-1}$	**0.5712**	0.1403	0.6221	-0.0157	0.2536	0.2082	0.8569		0.8738
$NPh2Inds_{t-2}$	**0.8331**	0.1875	0.2339	-0.0370	0.0909	-0.0192	0.8414		0.7646
$NPhase3_t$	**0.8446**	0.1584	0.0443	0.0517	0.0294	0.1280	0.7976	0.8331	0.7575
$NPhase3_{t-1}$	**0.5809**	0.1373	0.6798	0.0107	0.1421	0.0590	0.8457		0.8154
$NP3Inds_{t-2}$	**0.6957**	0.1776	0.1979	-0.0231	0.1403	0.1384	0.6128		0.7307
$NApprovals_t$	0.3188	0.2302	**0.4079**	-0.0323	0.0305	-0.1063	0.3877	0.6391	0.3866
$NApprovals_{t-1}$	0.5250	0.1419	**0.1286**	-0.0084	0.0508	0.0425	0.6289		0.6306
$NIndications_{t-2}$	0.2416	0.1519	**0.7532**	0.0061	0.0050	-0.0015	0.6634		0.6597
$Revenue_t$	0.1087	**0.8582**	0.1344	0.2168	-0.0089	0.0997	0.9502	0.9746	0.9543
$RDExpense_{t-2}$	0.1505	**0.8744**	0.1628	0.1699	0.0496	0.1073	0.9249		0.9620
$RDExpense_{t-3}$	0.1396	**0.8579**	0.1708	0.1970	0.0566	0.1497	0.9029		0.9709
$LicenseES_t$	0.2256	**0.6755**	0.1142	0.0768	0.1696	0.1184	0.5887	0.7450	0.5671
$LicESDealVal_t$	-0.0025	**0.6183**	-0.1031	0.1065	0.1026	0.0320	0.4953		0.7245
$LicESInds_{t-1}$	0.2909	**0.6147**	0.1383	0.0001	0.0526	0.2463	0.5822		0.6812
$LicPh12Deals_t$	0.1206	0.3158	0.0194	0.0582	0.0219	**0.6734**	0.5797	0.6271	0.3163
$LicPh12Deals_{t-1}$	0.2120	0.3680	0.1367	-0.1113	-0.0420	**0.1420**	0.2959		0.7927
$LicPh12DealVal_t$	0.2200	0.1750	0.0177	0.1101	0.0689	**0.7653**	0.6911		0.3782
$LicensePh3_t$	-0.0359	0.2099	-0.0118	**0.9476**	-0.0475	0.0642	0.9542	0.7665	0.4643
$LicPh3TAs_t$	-0.0435	0.2540	-0.0030	**0.9351**	-0.0505	0.1004	0.9566		0.4566
$LicPh3Inds_{t-1}$	0.0419	0.3402	-0.0015	**0.2052**	-0.0757	0.0633	0.3471		0.9843
Variance Explained by Each Factor	6.3885	4.2802	2.5972	2.0009	1.8017	1.4661			

Appendix B: Structural Equation Model for Non-Licensing Firms

Following the same rationale as in Sections 6.1 and 6.2, we develop a structural equation model for firms that do not engage in licensing, and we estimate the parameters of the model. Note that in this model, we do not have latent variables that represent firms' engagement in licensing.

B.1. Examining the Overall Measurement Model

As before, we first conduct the CFA of the measurement model and then estimate the parameters of the measurement model.

Factor analysis. Table 10 provides the standardized factor loadings. All factor loadings are at least 0.33 except $NPreclin_{t-1}$, $NPreclin_{t-2}$, $NPreclinInds_{t-3}$, $NPhase1_{t-2}$, $NPhase3_{t-2}$, and $NPh3Inds_{t-3}$. Similar to the SEM for the licensing firms, we keep these measured variables since they properly measure the R&D capabilities of the firms.

Reliability analysis. Table 10 lists the standardized Cronbach's alpha for each latent variable and the standardized alpha once a measured variable has been removed from the construct. The reliability measurement alphas for constructs *CapDiscover* and *CapFin* are at or above 0.6. However, the alphas for preclinical

Table 8 Measurement Model Estimates

Latent Variable	Measurement Variable	Estimates	t Statistic	90% Confidence Interval Limits		Error Variance	t Statistic	R-Square
				Lower	Upper			
$CapDiscover$	$NDiscovery_t$	1.00				1.77	13.25	0.563
	$NDiscovery_{t-1}$	0.85	17.15	0.77	0.93	0.93	12.64	0.638
	$NDiscovery_{t-2}$	0.54	15.62	0.49	0.60	0.57	13.36	0.542
$CapPreclin$	$NPreclin_t$	1.00				5.33	13.48	0.533
	$NPreclin_{t-1}$	0.74	17.43	0.67	0.81	2.55	13.27	0.564
	$NPreclinInds_{t-2}$	0.46	16.76	0.41	0.51	1.14	13.48	0.531
$CapPh1$	$NPhase1_t$	1.00				4.33	13.66	0.506
	$NPhase1_{t-1}$	1.45	21.78	1.34	1.56	2.11	11.30	0.815
	$NPh1Inds_{t-2}$	0.76	20.50	0.70	0.82	0.84	12.46	0.751
$CapPh2$	$NPhase2_t$	1.00				5.55	14.31	0.614
	$NPhase2_{t-1}$	1.04	20.28	0.95	1.12	3.56	13.73	0.727
	$NPh2Inds_{t-2}$	0.60	19.97	0.55	0.65	1.30	13.86	0.710
$CapPh3$	$NPhase3_t$	1.00				1.83	13.35	0.597
	$NPhase3_{t-1}$	0.87	19.03	0.79	0.94	0.87	12.37	0.702
	$NPhase3_{t-2}$	0.47	15.89	0.42	0.52	0.54	13.67	0.525
$CapProduct$	$NApprovals_t$	1.00				0.48	12.55	0.347
	$NApprovals_{t-1}$	0.81	9.86	0.68	0.95	0.27	12.19	0.388
	$NIndications_{t-2}$	0.75	10.06	0.63	0.87	0.21	11.96	0.410
$CapFin$	$Revenue_t$	1.00				17.71	7.38	0.951
	$RDExpense_{t-2}$	0.13	53.86	0.13	0.14	0.42	8.99	0.934
	$RDExpense_{t-3}$	0.12	48.00	0.12	0.13	0.55	10.78	0.904
$EngESLicense$	$LicenseES_t$	1.00				0.87	10.63	0.589
	$LicESDealVal_t$	0.11	10.71	0.09	0.13	0.03	12.94	0.318
	$LicESInds_{t-1}$	0.74	15.10	0.66	0.82	0.45	10.39	0.603
$EngPh12License$	$LicPh12Deals_t$	1.00				0.64	12.53	0.148
	$LicPh12Deals_{t-1}$	1.16	6.31	0.85	1.46	0.05	11.05	0.169
	$LicPh12DealVal_t$	0.34	10.13	0.28	0.39	0.73	12.05	0.212
$EngPh3License$	$LicensePh3_t$	1.00				0.40	11.76	0.250
	$LicPh3TAs_t$	0.95	35.95	0.91	1.00	0.30	11.16	0.288
	$LicPh3Inds_{t-1}$	0.81	7.69	0.64	0.99	0.21	11.00	0.297

capability ($CapPreclin$), R&D capability in ($CapPh1$), R&D capability in phase 2 ($CapPh2$), R&D capability in phase 3 ($CapPh3$), and product approval capability ($CapProduct$) are below 0.6. For the constructs $CapPreClin$, $CapPh1$, and $CapPh2$, the standardized Cronbach's alphas all decrease after we remove one measured variable (see the last column in Table 10). This observation implies that the measured variables for each of the above constructs are internally consistent. Removing any variable from each of the above constructs makes the construct less reliable. Note that the standardized alpha for the construct $CapPh3$ increases once $NPhase3_{t-1}$, $NPhase3_{t-2}$, or $NPh3Inds_{t-3}$ is removed and that the standardized alpha for the construct $CapProduct$ also increases once $Napprovals_{t-2} + Napprovals_{t-3}$ is removed. Nevertheless, we keep those three variables in their respective constructs because they properly represent firms' drug R&D capability in phase 3 and the drug approval stage.

Measurement model estimation. We estimate the parameters in the measurement model in which relations exist for all pairs of latent variables. Table 11 lists the fit indices in column (1). The chi-square value is 990, which is based on 271 degrees of freedom. The large chi-square is due to a large sample that contains 561 observations. Nonetheless, the ratio of the chi-square value to the degrees of freedom of the chi-square test is 3.65. The SRMR is 0.065, and the RMSEA is 0.069. The 95% confidence interval of the

Table 9 Structural Model Estimates

Latent Variable	Measurement Variable	Estimate	t Statistic	90% Confidence Interval Limits		Error Variance	t Statistic	R-Square
				Lower	Upper			
$CapDiscover$	$NDiscovery_t$	1.00				1.88	13.31	0.408
	$NDiscovery_{t-1}$	0.90	13.11	0.90	0.90	0.88	12.21	0.545
	$NDiscovery_{t-2}$	0.57	11.87	0.57	0.57	0.57	13.20	0.428
$CapPreclin$	$NPreclin_t$	1.00				3.90	13.67	13.665
	$NPreclin_{t-1}$	1.50	9.68	1.50	1.50	2.27	12.22	12.216
	$NPreclinInds_{t-2}$	0.90	9.32	0.90	0.90	1.12	12.87	12.873
$CapPh1$	$NPhase1_t$	1.00				3.50	12.52	0.716
	$NPhase1_{t-1}$	1.01	24.09	1.01	1.01	2.16	11.41	0.805
	$NPh1Inds_{t-2}$	0.54	23.29	0.54	0.54	0.73	11.88	0.777
$CapPh2$	$NPhase2_t$	1.00				4.52	12.93	0.677
	$NPhase2_{t-1}$	0.99	20.91	0.99	0.99	3.46	12.65	0.727
	$NPh2Inds_{t-2}$	0.59	21.63	0.59	0.59	1.05	12.41	0.758
$CapPh3$	$NPhase3_t$	1.00				1.70	12.93	0.620
	$NPhase3_{t-1}$	0.84	18.83	0.84	0.84	0.90	12.37	0.688
	$NPhase3_{t-2}$	0.48	16.38	0.48	0.48	0.50	13.24	0.556
$CapProduct$	$NApprovals_t$	1.00				0.46	12.01	0.364
	$NApprovals_{t-1}$	0.75	9.24	0.75	0.75	0.28	12.22	0.341
	$NIndications_{t-2}$	0.70	9.59	0.70	0.70	0.21	11.87	0.377
$CapFin$	$Revenue_t$	1.00				17.83	7.36	0.950
	$RDExpense_{t-2}$	0.13	53.55	0.13	0.13	0.41	8.87	0.934
	$RDExpense_{t-3}$	0.12	47.62	0.12	0.12	0.55	10.76	0.903
$EngESLicense$	$LicenseES_t$	1.00				0.81	9.65	0.567
	$LicESDealVal_t$	0.11	10.67	0.11	0.11	0.03	12.68	0.326
	$LicESInds_{t-1}$	0.71	14.23	0.71	0.71	0.48	10.42	0.611
$EngPh12License$	$LicPh12Deals_t$	1.00				0.55	10.93	0.267
	$LicPh12Deals_{t-1}$	1.05	6.79	1.05	1.05	0.05	11.19	0.264
	$LicPh12DealVal_t$	0.27	10.15	0.27	0.27	0.66	11.41	0.283
$EngPh3License$	$LicensePh3_t$	1.00				0.41	11.64	0.302
	$LicPh3TAs_t$	0.97	31.97	0.97	0.97	0.30	11.01	0.233
	$LicPh3Inds_{t-1}$	0.85	6.87	0.85	0.85	0.20	10.50	0.275

RMSEA is between 0.064 and 0.074. The SRMS, RMSEA, and the 95% confidence interval of the RMSEA indicate that the measurement model fits the covariance matrix of the measured variables well. The GFI is 0.9, which indicates a good model fit given the large sample of data we used. The Bentler comparative fit index, the Bentler-Bonett NFI, and the Bentler-Bonett non-normed index take on values of 0.89, 0.86, and 0.87, respectively. These incremental fit indices indicate a good fit of the measurement model.

Table 12 lists the factor loadings and t-statistics of the measured variables. All the measured variables have statistically significant relationships with their latent variables, their t-statistics exceed the critical value of the 0.1% significance level, and the measured variables are significantly entering the model.

B.2. Examining the Overall Structural Model

Table 11 lists the fit indices in column (2). The chi-square of the structural model is 992, which is based on 260 degrees of freedom. The ratio of the chi-square value to the degrees of freedom is 3.82, which corrects for model size. The SRMR is 0.066, and the RMSEA is 0.071. The 95% confidence interval of the RMSEA is between 0.066 and 0.076. The SRMR, RMSEA, and the 95% confidence interval of the RMSEA all indicate that the structural model fits the covariance matrix of the measured variables well, even though we restrict some path coefficients to be zero. The GFI is 0.88, which indicates a good model fit. The Bentler comparative

Table 10 Factor Analyses and Cronbachs Alpha Coefficients for Non-licensing Firms

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Community	Cronbach Alpha Coefficient	Cronbachs Alpha Coefficient with Deleted Variable
$NDiscovery_t$	-0.011	0.148	-0.087	**0.433**	0.220	0.563	0.450
$NDiscovery_{t-1}$	0.048	0.261	-0.073	**0.419**	0.283		0.503
$NDiscovery_{t-2}$	0.136	0.359	-0.117	**0.554**	0.476		0.431
$NPreclin_t$	-0.057	0.062	-0.068	**0.347**	0.140	0.244	0.231
$NPreclin_{t-1}$	-0.075	0.149	-0.033	**0.203**	0.071		0.169
$NPreclin_{t-2}$	-0.072	0.111	-0.020	**0.129**	0.038		0.225
$NPreclinInds_{t-3}$	-0.043	0.160	0.019	**0.117**	0.077		0.153
$NPhase1_t$	-0.192	**0.441**	0.004	-0.122	0.452	0.336	0.240
$NPhase1_{t-1}$	-0.155	**0.451**	0.065	-0.040	0.287		0.228
$NPhase1_{t-2}$	-0.079	**0.251**	0.005	0.112	0.129		0.304
$NPh1Inds_{t-3}$	-0.133	**0.423**	0.067	0.075	0.386		0.325
$NPhase2_t$	-0.136	**0.348**	-0.068	0.087	0.349	0.356	0.354
$NPhase2_{t-1}$	-0.147	**0.550**	0.079	-0.116	0.402		0.197
$NPhase2_{t-2}$	-0.126	**0.372**	0.066	0.011	0.278		0.270
$NPh2Inds_{t-3}$	-0.135	**0.416**	0.083	0.013	0.421		0.337
$NPhase3_t$	-0.097	**0.526**	0.057	-0.253	0.426	0.428	0.428
$NPhase3_{t-1}$	-0.144	**0.556**	0.176	-0.260	0.514		0.464
$NPhase3_{t-2}$	-0.104	**0.241**	0.070	-0.189	0.288		0.446
$NPh3Inds_{t-3}$	-0.108	**0.229**	0.087	-0.106	0.224		0.461
$Napprovals_t$	0.100	-0.054	**0.370**	-0.010	0.155	0.362	-0.005
$Napprovals_{t-1} + Napprovals_{t-2}$	0.162	-0.071	**0.835**	0.143	0.758		0.337
$Napprovals_{t-2} + Napprovals_{t-3}$	0.214	-0.093	**0.848**	0.143	0.804		0.435
$Revenue_t$	**0.896**	0.219	-0.085	-0.051	0.861	0.976	0.989
$RDExpense_{t-2}$	**0.949**	0.238	-0.097	-0.081	0.974		0.946
$RDExpense_{t-3}$	**0.940**	0.242	-0.090	-0.082	0.958		0.957

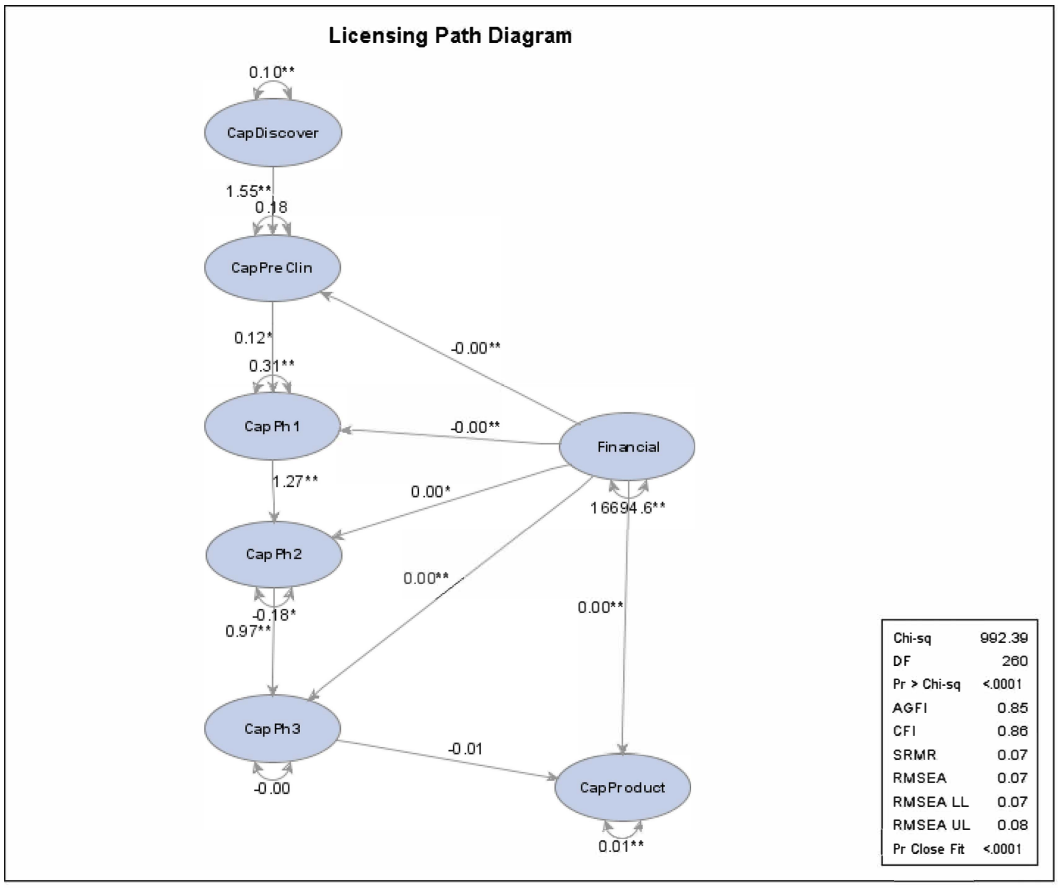
Table 11 Overall Fit Indices for Non-Licensing Firms

Index	Measurement Model (1)	Structural Model (2)
Chi-Square (χ^2)	990	992
Chi-Square (χ^2) Degree of Freedom	271	260
Standardized RMR (SRMR)	0.065	0.066
Goodness of Fit Index (GFI)	0.88	0.88
RMSEA	0.069	0.071
RMSEA Lower 90% Confidence Limit	0.064	0.066
RMSEA Upper 90% Confidence Limit	0.074	0.076
Bentler Comparative Fit Index	0.89	0.86
Bentler-Bonett NFI	0.86	0.82
Bentler-Bonett Non-normed Index	0.87	0.84

fit index, the Bentler-Bonett NFI, and the Bentler-Bonett non-normed index are 0.86, 0.82, and 0.84. These incremental indices all indicate a reasonably good model fit given our large sample size. Table 12 shows the path weight estimates. The factor loadings for the measured variables in the structural model are roughly the same as those in the measurement model, even though we reestimate them along with the relationship paths in the structural model. The consistency implies that the specification for the structural model is appropriate, which significantly represents the relationships among the latent variables. Therefore, the structural model replicates the covariance of the measured variables well. Figure 8 depicts the path coefficients that are different from zero at a 5% significance level, except the transition from the R&D capability in phase 3 ($CapPh3$) to the product approval capability ($CapProduct$).

Table 12 Measurement and Structural Model Estimates for Non-Licensing Firms

Measurement Variable	Latent Variable	Measurement Model		Structural Model	
		Estimate	t statistic	Estimate	t statistic
$NDiscovery_t$	$CapDiscover$	1.000		1.000	
$NDiscovery_{t-1}$		0.740	6.242	0.709	5.719
$NDiscovery_{t-2}$		0.951	6.962	1.212	5.838
$NPreclin_t$	$CapPreClin$	1.000		1.000	
$NPreclin_{t-1}$		0.379	4.362	0.304	3.129
$NPreclin_{t-2}$		0.269	3.910	0.283	3.560
$NPreclinInds_{t-3}$		0.255	4.636	0.175	2.900
$NPhase1_t$	$CapPh1$	1.000		1.000	
$NPhase1_{t-1}$		1.181	6.547	0.693	8.143
$NPhase1_{t-3}$		0.428	4.605	0.208	4.224
$NPh1Inds_{t-3}$		0.869	6.710	0.418	7.552
$NPhase2_{t-1}$	$CapPh2$	1.000		1.000	
$NPhase2_{t-2}$		1.297	6.808	1.270	6.990
$NPhase2_{t-3}$		0.613	5.768	0.550	5.541
$NPh2Inds_{t-3}$		0.716	6.596	0.537	6.023
$NPhase3_t$	$CapPh3$	1.000		1.000	
$NPhase3_{t-1}$		0.899	10.198	0.775	10.456
$NPhase3_{t-2}$		0.227	5.272	0.171	4.411
$NPh3Inds_{t-3}$		0.183	4.812	0.152	4.396
$Napprovals_t$	$CapProduct$	1.000		1.000	
$Napprovals_{t-1} + Napprovals_{t-2}$		1.319	7.771	1.150	8.772
$Napprovals_{t-2} + Napprovals_{t-3}$		2.746	5.525	1.846	6.913
$Revenue_t$	$CapFin$	1.000		1.000	
$RDExpense_{t-2}$		0.056	14.884	0.055	16.943
$RDExpense_{t-3}$		0.054	14.719	0.058	17.344



1

Figure 8 **Estimates of the Structural Equation Model for Non-licensing Firms.**