The Organization of Innovation: Incomplete Contracts & the Outsourcing Decision

Thomas Jungbauer[1] Sean Nicholson[2,3] June Pan[4]
Michael Waldman[1] **Lucy Xiaolu Wang**[5,6,7]

[1]Johnson Graduate School of Management, Cornell University
[2]Brooks School of Public Policy, Cornell University
[3]National Bureau of Economic Research
[4]Visa, Washington, DC
[5]Department of Resource Economics, University of Massachusetts Amherst
[6]Max Planck Institute for Innovation and Competition
[7]Canadian Centre for Health Economics

Productivity Lunch Seminar, NBER 2025.4.1

Research question:

Why do firms outsource research and development (R&D) for some products while conducting R&D in-house for similar ones?

Research question:

Why do firms outsource research and development (R&D) for some products while conducting R&D in-house for similar ones?

Potential explanation:

Superior expertise (past investment)

Research question:

Why do firms outsource research and development (R&D) for some products while conducting R&D in-house for similar ones?

Potential explanation:

Superior expertise (past investment)

But:

Unlikely explanation when required expertise is sufficiently similar

An Example

- In the 1990s and 2000, AstraZeneca, a research-based biopharmaceutical company, initiated 20 new anti-cancer and 5 new biotechnology drug developments.

An Example

- In the 1990s and 2000, AstraZeneca, a research-based biopharmaceutical company, initiated 20 new anti-cancer and 5 new biotechnology drug developments.
- While it outsourced development of all biotechnology drugs, it developed all anti-cancer drugs in-house.
- At the time, AstraZeneca did not hold a single patent in the biotechnology class but had filed for 9 patents for anti-cancer drugs, including the well-known and highly successful compounds "Cosudex" and "Iressa."

Motivation

- We postulate that firms are systematically more reluctant to relinquish control over R&D, the more successfully they already operate in the same product category.

Motivation

- We postulate that firms are systematically more reluctant to relinquish control over R&D, the more successfully they already operate in the same product category.
- Successful products are threatened by innovation.
- Naturally, a firm will want to deter the introduction of close substitutes to a profitable product.

Motivation

- We postulate that firms are systematically more reluctant to relinquish control over R&D, the more successfully they already operate in the same product category.
- Successful products are threatened by innovation.
- Naturally, a firm will want to deter the introduction of close substitutes to a profitable product.
- These new products may be introduced by competitors (Cunningham et al., 2021), or the firm itself (cannibalization).

Motivation

- We postulate that firms are systematically more reluctant to relinquish control over R&D, the more successfully they already operate in the same product category.
- Successful products are threatened by innovation.
- Naturally, a firm will want to deter the introduction of close substitutes to a profitable product.
- These new products may be introduced by competitors (Cunningham et al., 2021), or the firm itself (cannibalization).
- The property rights theory of the firm suggest that contracts are inherently incomplete and that in-house development increases the level control over R&D.

Preview

We build a model in which a firm, the originator, develops a new product and decides whether to outsource the remaining R&D.

Preview

We build a model in which a firm, the originator, develops a new product and decides whether to outsource the remaining R&D.

- The originator owns an existing patented product in the same product category.

Preview

We build a model in which a firm, the originator, develops a new product and decides whether to outsource the remaining R&D.

- The originator owns an existing patented product in the same product category.
- The firm that conducts R&D is unable to choose the exact location of the new product in product space, but chooses a mean location and (costly) precision.

Preview

We build a model in which a firm, the originator, develops a new product and decides whether to outsource the remaining R&D.

- The originator owns an existing patented product in the same product category.
- The firm that conducts R&D is unable to choose the exact location of the new product in product space, but chooses a mean location and (costly) precision.
- There is a competitive pool of licensees that may be more efficient than the originator.

Preview

We build a model in which a firm, the originator, develops a new product and decides whether to outsource the remaining R&D.

- The originator owns an existing patented product in the same product category.
- The firm that conducts R&D is unable to choose the exact location of the new product in product space, but chooses a mean location and (costly) precision.
- There is a competitive pool of licensees that may be more efficient than the originator.
- Product location is not contractible.

Theoretical Results = Testable Predictions

Existing product(s): A firm is less likely to outsource R&D for a new product if it has existing products in the same product category.

Theoretical Results = Testable Predictions

Existing product(s): A firm is less likely to outsource R&D for a new product if it has existing products in the same product category.

Patent(s) length: A firm is less likely to outsource R&D for a new product, the longer the patent(s) of its existing product(s).

Theoretical Results = Testable Predictions

Existing product(s): A firm is less likely to outsource R&D for a new product if it has existing products in the same product category.

Patent(s) length: A firm is less likely to outsource R&D for a new product, the longer the patent(s) of its existing product(s).

Pivots: A firm's new product will be closer to the optimal location if developed in-house, when related patents are yet to expire.

Theoretical Results = Testable Predictions

Existing product(s): A firm is less likely to outsource R&D for a new product if it has existing products in the same product category.

Patent(s) length: A firm is less likely to outsource R&D for a new product, the longer the patent(s) of its existing product(s).

Pivots: A firm's new product will be closer to the optimal location if developed in-house, when related patents are yet to expire.

Competition: A firm is less likely to outsource R&D for a new product, the higher the market share of its existing product(s).

Empirical Analysis

- We test these results using comprehensive data from the pharmaceutical industry, i.e.,
  - drug development data from the Pharmaprojects dataset,
  - and drug sales data from IMS dataset.

Empirical Analysis

- We test these results using comprehensive data from the pharmaceutical industry, i.e.,
  - drug development data from the Pharmaprojects dataset,
  - and drug sales data from IMS dataset.
- We run reduced-form Logit regressions to predict the probability of outsourcing R&D for a new drug development project given
  - patent existence, patent length, and market share with different sets of controls.

Empirical Analysis

- We test these results using comprehensive data from the pharmaceutical industry, i.e.,
  - drug development data from the Pharmaprojects dataset,
  - and drug sales data from IMS dataset.
- We run reduced-form Logit regressions to predict the probability of outsourcing R&D for a new drug development project given
  - patent existence, patent length, and market share with different sets of controls.
- OLS regressions to link the number of R&D pivots with in-house development status and observable controls.

Empirical Analysis

- We test these results using comprehensive data from the pharmaceutical industry, i.e.,
  - drug development data from the Pharmaprojects dataset,
  - and drug sales data from IMS dataset.
- We run reduced-form Logit regressions to predict the probability of outsourcing R&D for a new drug development project given
  - patent existence, patent length, and market share with different sets of controls.
- OLS regressions to link the number of R&D pivots with in-house development status and observable controls.
- Our empirical results strongly support the validity of all four testable predictions.

Empirical Analysis

- We test these results using comprehensive data from the pharmaceutical industry, i.e.,
  - drug development data from the Pharmaprojects dataset,
  - and drug sales data from IMS dataset.
- We run reduced-form Logit regressions to predict the probability of outsourcing R&D for a new drug development project given
  - patent existence, patent length, and market share with different sets of controls.
- OLS regressions to link the number of R&D pivots with in-house development status and observable controls.
- Our empirical results strongly support the validity of all four testable predictions.
- Note that our theoretical results are not confined to the pharmaceutical industry.

Related Literature

- **Property Rights**: Grossman and Hart (1986), Hart and Moore (1990), Aghion and Tirole (1994), Lerner and Merges (1998), Gibbons (2005), Lafontaine and Slade (2007);
- **Vertical Integration and Competition**: Tucker and Wilder (1977), Levy (1985), Balakrishnan and Wernerfelt (1986), Galdon-Sanchez et al. (2015), Gil and Ruzzier (2018);
- **Cannibalization and Obsolescence**: Moorthy and Png (1992), Waldman (1993, 1996), Choi (1994), Nahm (2004), Igami (2017);
- **Pharmaceutical Industry**: DiMasi et al. (2003), Azoulay (2004), Nicholson et al. (2005), Williams (2013), Budish et al. (2015), Krieger et al. (2017), Lakdawalla (2018);

R&D and Production

The existing product (EP)

- The originator's marginal cost per unit is c_1.
- Generic producers produce at c_1 after patent expiration at t_E.

R&D and Production

The existing product (EP)

- The originator's marginal cost per unit is c_1.
- Generic producers produce at c_1 after patent expiration at t_E.

The new product (NP)

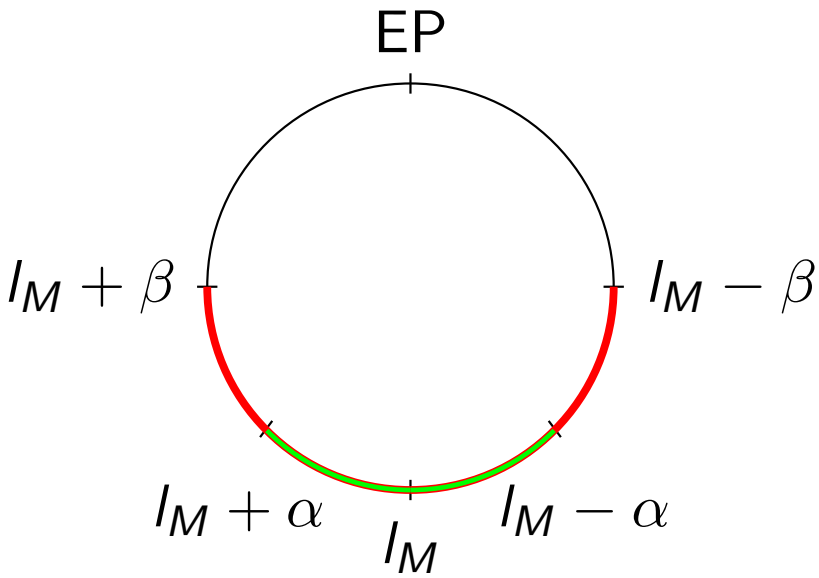
- The originator faces a fixed cost of development F, which is a random draw of $f(\cdot)$ over $[F_{min}, \infty)$.
- A licensee faces a fixed cost F_L, $F_{min} \leq F_L < \infty$, $\Delta = F - F_L$.
- The developer incurs a per unit production cost of c_2, while a firm that did not develop the product incurs c_2^+, $c_2^+ > c_2$.
- The patent expires at $t = T$.

Product Space

The new product's location is $l = l_M + \varepsilon$.

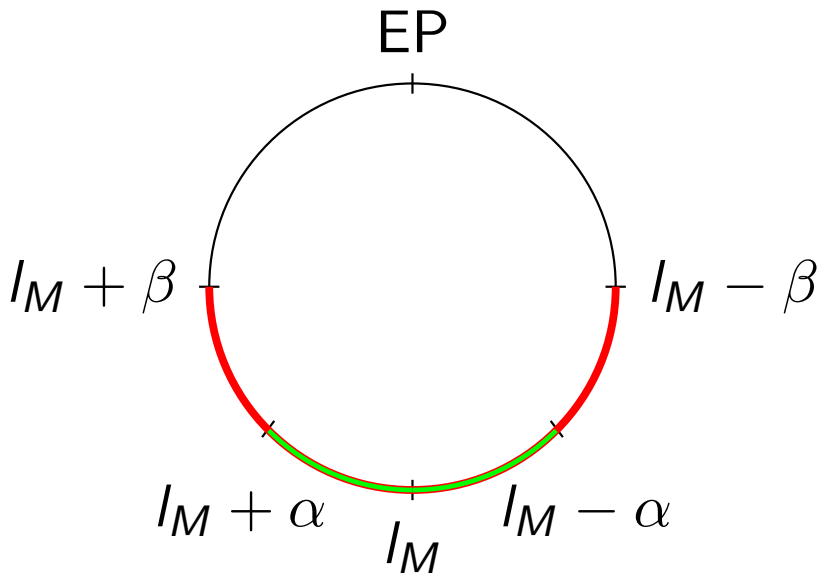
Product Space

The new product's location is $l = l_M + \varepsilon$.



Product Space

The new product's location is $l = l_M + \varepsilon$.



The investment level k determines whether ε is drawn from $U[-\alpha, \alpha]$ (with probability $p(k)$) or $U[-\beta, \beta]$ (with $1 - p(k)$).

Consumers

- Consumers are uniformly located around the unit circle.
- Each consumer decides for one or the other product.

Consumers

- Consumers are uniformly located around the unit circle.
- Each consumer decides for one or the other product.
- The value a consumer draws from consuming the q-th unit of a product is given by

$$V^{+} - vq. \tag{1}$$

Consumers

- Consumers are uniformly located around the unit circle.
- Each consumer decides for one or the other product.
- The value a consumer draws from consuming the q-th unit of a product is given by

$$V^+ - vq. \tag{1}$$

- Therefore, each consumer chooses Q to maximize

$$\int_0^Q (V^+ - vq)dq - (P + ds)Q \tag{2}$$

given her distance s from the product and its price P.

Contracting

If the originator decides to outsource production to licensee, they sign a contract. The contract specifies for each period who

Contracting

If the originator decides to outsource production to licensee, they sign a contract. The contract specifies for each period who

- produces the product,
- sells the product (collects revenue),
- sets the prize,
- and a payment from the originator to the licensee.

Contracting

If the originator decides to outsource production to licensee, they sign a contract. The contract specifies for each period who

- produces the product,
- sells the product (collects revenue),
- sets the prize,
- and a payment from the originator to the licensee.

We assume that

- the payment can depend on the period's quantity,
- the contract cannot depend on location, and
- that the contract is renegotiation-proof.

Time line

- **t=1**
  - Originator chooses price p_E^1, and consumers purchase
  - F is realized and publicly revealed
  - Originator chooses whether to outsource development of NP
  - If yes, licensees make take-it or leave-it contract offers

Time line

- **t=1**
  - Originator chooses price p_E^1, and consumers purchase
  - F is realized and publicly revealed
  - Originator chooses whether to outsource development of NP
  - If yes, licensees make take-it or leave-it contract offers
- **t=2**
  - Originator chooses p_E^2, and consumers purchase
  - Developer chooses l_M and k, ε is realized
  - NP's location is publicly known, but not verifiable

Time line

- **t=1**
  - Originator chooses price p_E^1, and consumers purchase
  - F is realized and publicly revealed
  - Originator chooses whether to outsource development of NP
  - If yes, licensees make take-it or leave-it contract offers
- **t=2**
  - Originator chooses p_E^2, and consumers purchase
  - Developer chooses l_M and k, ε is realized
  - NP's location is publicly known, but not verifiable
- **t=3**
  - NP is introduced
  - Originator chooses p_E^3, originator or licensee chooses p_N^3
  - Consumers purchase
- **...**

The Equilibrium Contract

Note that throughout the paper we focus on Subgame Perfect Nash Equilibrium (SPNE).

Lemma

Suppose the originator chooses outsourcing.

The Equilibrium Contract

Note that throughout the paper we focus on Subgame Perfect Nash Equilibrium (SPNE).

Lemma

Suppose the originator chooses outsourcing.

(a) Then, for $t \leq t_E$, the contract assigns production to the licensee, but sales and pricing to the originator.

The Equilibrium Contract

Note that throughout the paper we focus on Subgame Perfect Nash Equilibrium (SPNE).

Lemma

Suppose the originator chooses outsourcing.

(a) Then, for $t \leq t_E$, the contract assigns production to the licensee, but sales and pricing to the originator.

(b) For $t > t_E$, the contract assigns production, sales, and pricing to the licensee.

The Equilibrium Contract

Note that throughout the paper we focus on Subgame Perfect Nash Equilibrium (SPNE).

Lemma

Suppose the originator chooses outsourcing.

(a) Then, for $t \leq t_E$, the contract assigns production to the licensee, but sales and pricing to the originator.

(b) For $t > t_E$, the contract assigns production, sales, and pricing to the licensee.

(c) The payments from the originator to the licensee sum to the fixed amount that guarantees the licensee zero ex ante profits.

Location & Precision Choice

Let I refer to in-house while O refers to outsourcing.

Lemma

Let $L(j, t_E)$ and $K(j, t_E)$ be firm j's location and precision choice in equilibrium given t_E. Then,

Location & Precision Choice

Let I refer to in-house while O refers to outsourcing.

Lemma

Let $L(j, t_E)$ and $K(j, t_E)$ be firm j's location and precision choice in equilibrium given t_E. Then,

(a) $L(I, t_E) = L(O, t_E) = \frac{1}{2}\ \forall t_E > 1$,

Location & Precision Choice

Let I refer to in-house while O refers to outsourcing.

Lemma

Let $L(j, t_E)$ and $K(j, t_E)$ be firm j's location and precision choice in equilibrium given t_E. Then,

(a) $L(I, t_E) = L(O, t_E) = \frac{1}{2}\ \forall t_E > 1$,

(b) $K(I, 1) = K(I, 2) = K(O, 1) = K(O, 2)$,

Location & Precision Choice

Let I refer to in-house while O refers to outsourcing.

Lemma

Let $L(j, t_E)$ and $K(j, t_E)$ be firm j's location and precision choice in equilibrium given t_E. Then,

(a) $L(I, t_E) = L(O, t_E) = \frac{1}{2}\ \forall t_E > 1$,

(b) $K(I, 1) = K(I, 2) = K(O, 1) = K(O, 2)$,

(c) $K(I, t_E) > K(O, t_E)\ \forall t_E > 2$ *and* $K(O, T) = 0$,

Location & Precision Choice

Let I refer to in-house while O refers to outsourcing.

Lemma

Let $L(j, t_E)$ and $K(j, t_E)$ be firm j's location and precision choice in equilibrium given t_E. Then,

(a) $L(I, t_E) = L(O, t_E) = \frac{1}{2}\ \forall t_E > 1$,

(b) $K(I, 1) = K(I, 2) = K(O, 1) = K(O, 2)$,

(c) $K(I, t_E) > K(O, t_E)\ \forall t_E > 2$ *and* $K(O, T) = 0$,

(d) $K(I, T) > K(I, T-1) > ... > K(I, 2) = K(I, 1)$,

Location & Precision Choice

Let I refer to in-house while O refers to outsourcing.

Lemma

Let $L(j, t_E)$ and $K(j, t_E)$ be firm j's location and precision choice in equilibrium given t_E. Then,

(a) $L(I, t_E) = L(O, t_E) = \frac{1}{2}\ \forall t_E > 1$,

(b) $K(I, 1) = K(I, 2) = K(O, 1) = K(O, 2)$,

(c) $K(I, t_E) > K(O, t_E)\ \forall t_E > 2$ *and* $K(O, T) = 0$,

(d) $K(I, T) > K(I, T-1) > ... > K(I, 2) = K(I, 1)$,

(e) $K(O, 1) = K(O, 2) > K(O, 3) > ... > K(O, T) = 0$.

Patent Existence and Length

Recall that $\Delta = F - F_L$.

.

Patent Existence and Length

Recall that $\Delta = F - F_L$.

.

Proposition

(a) There exists a value $\Delta^* \geq 0$ such that the originator chooses outsourcing if $\Delta > \Delta^*$ and chooses in-house development otherwise.

Patent Existence and Length

Recall that $\Delta = F - F_L$.

.

Proposition

(a) There exists a value $\Delta^* \geq 0$ such that the originator chooses outsourcing if $\Delta > \Delta^*$ and chooses in-house development otherwise.

(b) Δ^* is an increasing function of t_E, and $\Delta^* = 0$ if and only if $t_E = 1$ or $t_E = 2$.

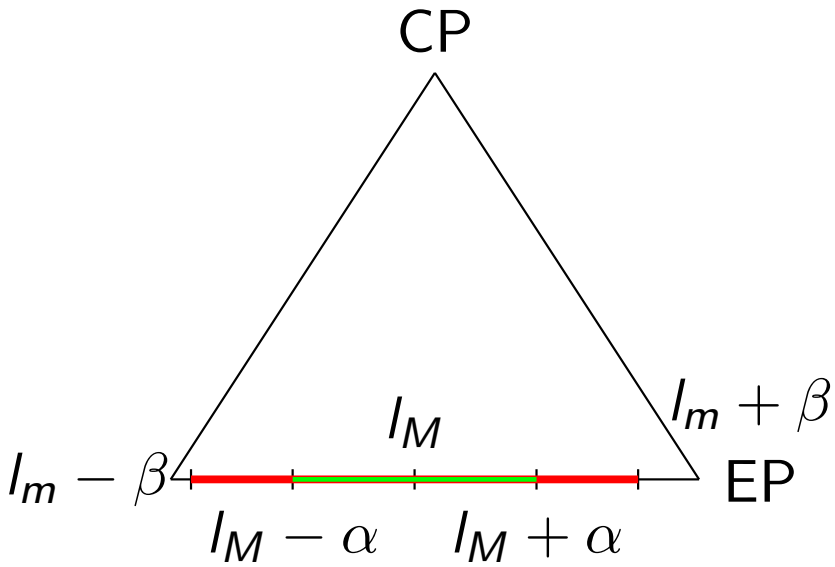
(c) Equilibrium behavior is essentially unique.

Competition

Consider now a simplified 2-period model with competition, in which the existing product (EP) and a competitor's product (CP) are located on two corners of the equilateral unit triangle.

Competition

Consider now a simplified 2-period model with competition, in which the existing product (EP) and a competitor's product (CP) are located on two corners of the equilateral unit triangle.



Competition and Market Share

- EP and CP provide consumers with different levels of utility.
- Uniformly distributed consumers buy only one product.
- The originator introduces a new product in the second period on the side opposite to the competitor.

Competition and Market Share

- EP and CP provide consumers with different levels of utility.
- Uniformly distributed consumers buy only one product.
- The originator introduces a new product in the second period on the side opposite to the competitor.

Proposition

In the unique equilibrium of the 2-period innovation game with competition, the originator is less likely to outsource the higher its market share if it did not introduce a new product.

Testable Predictions

Testable Prediction 1: A firm has a lower probability to outsource if it sells an existing product in the same product class.

Testable Predictions

Testable Prediction 1: A firm has a lower probability to outsource if it sells an existing product in the same product class.

Testable Prediction 2: A firm has a lower probability to outsource the longer the patent(s) of its existing product(s).

Testable Predictions

Testable Prediction 1: A firm has a lower probability to outsource if it sells an existing product in the same product class.

Testable Prediction 2: A firm has a lower probability to outsource the longer the patent(s) of its existing product(s).

Testable Prediction 3: A firm's new product is less likely to pivot from the primary therapeutic areas when developed in-house, esp. given the existence of related patents that have not expired.

Testable Predictions

Testable Prediction 1: A firm has a lower probability to outsource if it sells an existing product in the same product class.

Testable Prediction 2: A firm has a lower probability to outsource the longer the patent(s) of its existing product(s).

Testable Prediction 3: A firm's new product is less likely to pivot from the primary therapeutic areas when developed in-house, esp. given the existence of related patents that have not expired.

Testable Prediction 4: A firm has a lower probability of outsourcing the higher the predicted market share of its existing product(s) when the new product is introduced.

Data and Sample

The Pharmaprojects dataset: drug development data 1989-2004

- Assembled by the company Informa
- Contains information about drug development and new pharmaceutical projects by publicly traded firms worldwide

Data and Sample

The Pharmaprojects dataset: drug development data 1989-2004

- Assembled by the company Informa
- Contains information about drug development and new pharmaceutical projects by publicly traded firms worldwide

The IMS dataset: all drug sales in the US 1992-2004

- Merged to developed drugs from the Pharmaprojects dataset
- Calculated market shares with and without imputation

Data and Sample

The Pharmaprojects dataset: drug development data 1989-2004

- Assembled by the company Informa
- Contains information about drug development and new pharmaceutical projects by publicly traded firms worldwide

The IMS dataset: all drug sales in the US 1992-2004

- Merged to developed drugs from the Pharmaprojects dataset
- Calculated market shares with and without imputation

Therapeutic class crosswalk with ATC and firm-level info collection

- Constructed supporting data from various public sources

Data and Sample

The Pharmaprojects dataset: drug development data 1989-2004

- Assembled by the company Informa
- Contains information about drug development and new pharmaceutical projects by publicly traded firms worldwide

The IMS dataset: all drug sales in the US 1992-2004

- Merged to developed drugs from the Pharmaprojects dataset
- Calculated market shares with and without imputation

Therapeutic class crosswalk with ATC and firm-level info collection

- Constructed supporting data from various public sources

Main analysis: logit regressions and OLS regressions

- 109,115 compound-year observations 1989-2004

▸ summary of drug dev phases ▸ therapeutic classification ▸ firm types: overall, in-house, and outsource

Definition of Constructed Variables

Variable	Description
In-house	Indicator equals 1 if compound is never contracted out by the originating firm or if its earliest Development Contract was made after the start of Phase III trials.
Existence of Patents	
EOP1	Indicator equals 1 if at least one other compound in the same therapeutic class and same firm is patented.
EOP2	Number of other patented compounds in the same therapeutic class and same firm.
Length of Patents	
LOP1	Length of the longest patent among compounds in the same therapeutic class and same firm.
LOP2	Sum of the patent lengths among compounds in the same therapeutic class and same firm.
Other Variables	
Experience	Cumulative count of compound-year observations within a firm for a therapeutic class corresponding to the compound of interest.
Scope	Sum of the squares of the percentage of compounds being developed for each therapeutic class within a firm in a given year.
PDM	Number of patented drugs on the market in the same therapeutic class but not the same firm as the compound of interest.
TDM	Total number of drugs on the market in the same therapeutic class as the compound of interest.
MSP	Market share based on sales for existing patented drugs in the same class and same firm as the compound of interest.

Descriptive Statistics on In-house Dev and Patent Profile

Number of compounds					11,493
Number of firms					532
Years covered					1989-2004
	Outsourced Mean (SD)	**In-house** Mean (SD)	**Overall** Mean (SD)	**Min**	**Max**
Level of Observation: Compound-Year (109,115)					
In-house	0 (0)	1 (0)	0.785 (0.411)	0	1
Existence of Patents					
EOP1	0.586 (0.493)	0.766 (0.424)	0.727 (0.445)	0	1
EOP2	4.370 (8.550)	11.030 (14.007)	9.597 (13.312)	0	64
Length of Patents					
LOP1	10.490 (7.881)	12.967 (7.223)	12.434 (7.440)	0	20
LOP2	56.827 (103.235)	135.479 (169.847)	118.553 (161.177)	0	884

▸ summary stats for control variables

Logit Models of In-house Dev: Existence of Patents (EOP)

	(1) In-house	(2) In-house	(3) In-house	(4) In-house
EOP1	0.899***	0.505***		
	(0.0550)	(0.0600)		
EOP2			0.0632***	0.0595***
			(0.00430)	(0.00531)
Phase I	-0.649***	-0.674***	-0.622***	-0.640***
	(0.0799)	(0.0801)	(0.0796)	(0.0799)
Phase II	-1.031***	-1.111***	-1.017***	-1.034***
	(0.0692)	(0.0704)	(0.0711)	(0.0712)
Phase III	-1.425***	-1.542***	-1.436***	-1.462***
	(0.0906)	(0.0931)	(0.0943)	(0.0936)
Launched	-2.018***	-2.174***	-1.996***	-2.021***
	(0.126)	(0.129)	(0.127)	(0.126)
Experience		0.00151***		-0.000167
		(0.000142)		(0.000132)
Scope		-1.467***		-1.360***
		(0.150)		(0.148)
Observations	109,115	109,115	109,115	109,115

Logit Models of In-house Dev: Length of Patents (LOP)

	(1) In-house	(2) In-house	(3) In-house	(4) In-house
LOP1	0.0518***	0.0271***		
	(0.00333)	(0.00356)		
LOP2			0.00525***	0.00407***
			(0.000353)	(0.000334)
Phase I	-0.657***	-0.680***	-0.631***	-0.652***
	(0.0792)	(0.0797)	(0.0798)	(0.0800)
Phase II	-1.043***	-1.120***	-1.022***	-1.058***
	(0.0693)	(0.0706)	(0.0711)	(0.0713)
Phase III	-1.466***	-1.569***	-1.452***	-1.501***
	(0.0902)	(0.0930)	(0.0942)	(0.0942)
Launched	-2.009***	-2.174***	-1.995***	-2.065***
	(0.123)	(0.128)	(0.125)	(0.126)
Experience		0.00157***		0.000439***
		(0.000144)		(8.58e-05)
Scope		-1.602***		-1.333***
		(0.149)		(0.149)
Observations	109,115	109,115	109,115	109,115

Number of Pivots by In-house Development Status

	(1) # pivot	(2) # pivot	(3) # pivot
In-house	-0.249***	-0.216***	-0.217***
	(0.0397)	(0.0381)	(0.0381)
In-house x Patent	-0.108**	-0.117**	-0.113**
	(0.0520)	(0.0503)	(0.0503)
Patent	0.111**	0.116**	0.117**
	(0.0481)	(0.0463)	(0.0464)
Constant	0.726***	0.183***	-0.136
	(0.0362)	(0.0359)	(0.108)
Therapeutic class FE		Yes	Yes
Firm type FE			Yes
Observations	11,493	11,493	11,493

Notes: Each unit is a drug project. Dependent variable is the number of pivots, calculated as the total number of other additional therapeutic classes (second level ATC) a drug is tested or intended to be tested for, in addition to the main therapeutic class the project is seeking approval for or being approved. Column 1 reports the regression with in-house, patent existence, and the interaction term. Column 2 further includes a set of primary therapeutic class fixed effects. Column 3 further adds firm type fixed effects to the specification in Column 2.

Logit Models of In-house Dev: Patents with Interactions

	(1) In-house	(2) In-house	(3) In-house	(4) In-house
EOP1	1.164***			
	(0.140)			
EOP1 × PDM	-0.0251***			
	(0.00447)			
EOP2		0.129***		
		(0.0170)		
EOP2 × PDM		-0.00239***		
		(0.000484)		
LOP1			0.0588***	
			(0.00809)	
LOP1 × PDM			-0.00120***	
			(0.000255)	
LOP2				0.00959***
				(0.00125)
LOP2 × PDM				-0.000186***
				(3.60e-05)
Observations	109,115	109,115	109,115	109,115

Current Market Share

	(1) In-house	(2) In-house	(3) In-house	(4) In-house	(5) In-house
Panel A					
Current MSP	0.0377***	0.0330***	0.0248***	0.0333***	0.0234***
	(0.00578)	(0.00565)	(0.00559)	(0.00568)	(0.00543)
EOP1		0.408***			
		(0.0610)			
EOP2			0.0492***		
			(0.00515)		
LOP1				0.0207***	
				(0.00371)	
LOP2					0.00341***
					(0.000319)
Observations	101,586	101,586	101,586	101,586	101,586

Notes: Dependent variable is one if a compound is developed in-house, and zero otherwise. All specifications include experience, scope, development phase indicators, and a full set of therapeutic category and year indicators. We control for current market share based on sales for existing patented drugs in the same class and same firm as the compound of interest. Panel A includes a larger sample due to imputing missing MSP values as zero for firms without positive sales in a given therapeutic area, assuming their absence indicates a lack of sales activity.

▸ Results using alternative market share measure

Future Market Share

	(1) In-house	(2) In-house	(3) In-house	(4) In-house	(5) In-house
Panel A					
Future MSP	0.0431***	0.0351***	0.0322***	0.0364***	0.0298***
	(0.00849)	(0.00806)	(0.00852)	(0.00818)	(0.00833)
EOP1		0.650***			
		(0.0916)			
EOP2			0.0731***		
			(0.0116)		
LOP1				0.0276***	
				(0.00495)	
LOP2					0.00455***
					(0.000670)
Observations	14,468	14,468	14,468	14,468	14,468

Notes: Dependent variable is one if a compound is developed in-house, and zero otherwise. All specifications include experience, scope, development phase indicators, and a full set of therapeutic category and year indicators. We control for future market share based on future drug sales in the same therapeutic category by the same firm, assuming perfect fullsight and average development length as in DiMasi et al. (2003). Panel A includes a larger sample due to imputing missing MSP values as zero for firms without positive sales in a given therapeutic area.

▸ Results using alternative future market share measure

Robustness Checks

- Outsourcing definition:
  - What if we ignore phases I to II and define outsourcing/in-house development based on phase I?
  - All main results are qualitatively robust to this alternative defn.

▸ alternative in-house defn (before phase II)

Robustness Checks

- Outsourcing definition:
  - What if we ignore phases I to II and define outsourcing/in-house development based on phase I?
  - All main results are qualitatively robust to this alternative defn.

    ▸ alternative in-house defn (before phase II)
- Product categorization:
  - Are our results driven by too broad therapeutic categories (e.g., therapeutic class of "cardiovascular system")?
  - Results are robust to narrow categories (e.g., antihypertensive, diuretics, peripheral vasodilators, vasoprotectives, agents act on reninangiotensin system, and lipid modifying agents). ▸ finer ATC
- Firm-level characteristics:
  - Are our results driven by firm (type) specific characteristics (e.g., pharma vs NPOs, portfolio features, firm-level factors)?
  - Results are robust to various fixed effects at the firm (type) level or observable controls ▸ w/ firm type FE ▸ excl. CROs/CDMOs ▸ biopharma firms only ▸ w/ firm portfolio ctrls ▸ w/ firm FE

Robustness Checks (continue)

- Development speed test:
  - Does in-house development speed up the development process? It does seem to be the case. ▸ development speed test
  - Future research along these lines on optimal timing rather than solely development speed seems to us of particular interest.

Robustness Checks (continue)

- Development speed test:
  - Does in-house development speed up the development process? It does seem to be the case. ▸ development speed test
  - Future research along these lines on optimal timing rather than solely development speed seems to us of particular interest.
- Subsample analysis:
  - Exclude later development stages: exclude observations that occur after the end of phase II. ▸ subsample by end of phase II
  - Exclude compound years with ownership changes (e.g., due to mergers & acquisitions). ▸ subsample w/o ownership changes
  - Exclude observations that are potentially right-censored if a drug began a phase within the 95% completion threshold but has not yet completed it. ▸ adjusted for potential censoring

Alternative Theories

The prior literature and common sense support two alternative theories that could possibly explain the effects we observe:

Alternative Theories

The prior literature and common sense support two alternative theories that could possibly explain the effects we observe:

- Research incentives vs. financing cost: à la Aghion and Tirole (1994)
  - Integrated structure chosen when research incentives are more important than financial considerations

Alternative Theories

The prior literature and common sense support two alternative theories that could possibly explain the effects we observe:

- Research incentives vs. financing cost: à la Aghion and Tirole (1994)
  - Integrated structure chosen when research incentives are more important than financial considerations
  - Focus on successful development, not location (our case)
  - Does not account for why patents in the same category should be particularly important, nor the frequency of pivots in R&D

Alternative Theories

The prior literature and common sense support two alternative theories that could possibly explain the effects we observe:

- Research incentives vs. financing cost: à la Aghion and Tirole (1994)
  - Integrated structure chosen when research incentives are more important than financial considerations
  - Focus on successful development, not location (our case)
  - Does not account for why patents in the same category should be particularly important, nor the frequency of pivots in R&D
- Learning curve argument:
  - Expertise developed in prior R&D investments can lower costs
  - But does not explain our results on pivots, esp. when originator owns an existing patent product in the same therapeutic class
- Others: learning-by-doing, data/knowledge intensive activities

  (not good matches to our results regarding patents/pivots)

Takeaways

- Limiting cannibalization of existing products is an important factor in the decision whether to outsource R&D.
- Outsourcing diminishes managerial control.
- We build a theoretical model that predicts that a firm is more likely to outsource R&D
  - if it sells existing products in the same class,
  - the longer the patents of these products,
  - and the larger their market share.
  - In-house R&D projects pivot less from main therapeutic areas.
- An empirical analysis of the pharmaceutical industry confirms our testable predictions.
- Future research: How does competition in innovation affect these findings and vice versa? Do the predictions match evidence in other industries?


thank you
danke
謝謝
рахмат
ngiyabonga
teşekkür ederim
gracias
спасибо
Баярлалаа
faafetai lava
vinaka
спасибі
blagodaram
dank je
misaotra
matondo
paldies
grazzi
tapadh leat
хвала
asante
manana
obrigada
tenki
murakoze
chokrane
mochchakkeram
djiere dieuf
tau
mamnun
дякую
bedankt
dziękuję
hvala
mauruuru
kiitos
dankie
dhanyavad
nanni
nandri
bayarlalaa
grâçie
köszönöm
enkosi
dankon
akun
aciū
go raibh maith agat
obrigado
sukriya
kop khun krap
sagolun
sobodi
dèkuji
chnorakaloutioun
gratias ago
gràcies
sulpáy
taiku
arigatô
takk
dakujem
trugarez
grazie
ありがとう
tanemirt
rahmet
najis tuke
terima kasih
didi madloba
kam sah hamnida
rahmat
dhanyavadagalu
shukriya
diolch
merci
мерси
mèsi
merce
তোমাকে ধন্যবাদ
감사합니다
xièxie
ευχαριστώ

Aghion, P. and J. Tirole (1994). The Management of Innovation. *Quarterly Journal of Economics 109*, 1185–1209.

Azoulay, P. (2004). Capturing Knowledge Within and Across Firm Boundaries: Evidence from Clinical Development. *American Economic Review 94*, 1591–1612.

Balakrishnan, S. and B. Wernerfelt (1986). Technical Change, Competition and Vertical Integration. *Strategic Management Journal 7*, 347–359.

Budish, E., B. N. Roin, and H. Williams (2015). Do Firms Underinvest in Long-Term Research? Evidence from Cancer Clinical Trials. *American Economic Review 105*, 2044–2085.

Choi, J. P. (1994). Network Externality, Compatibility Choice, and Planned Obsolescence. *The Journal of Industrial Economics 42*, 167–182.

Cunningham, C., F. Ederer, and S. Ma (2021). Killer Acquisitions. *Journal of Political Economy 129*, 649–702.

DiMasi, J. A., R. W. Hansen, and H. G. Grabowski (2003). The Price of Innovation: New Estimates of Drug Development Costs. *Journal of Health Economics 22*, 151–185.

Galdon-Sanchez, J. E., R. Gil, and A. Bayo-Mariones (2015). Outsourcing of Peripheral Services: Evidence from Spanish Manufacturing Plant-Level Data. *European Economic Review 78*, 328–344.

Gibbons, R. (2005). Four Formal(izable) Theories of the Firm? *Journal of Economic Behavior & Organization 58*, 200–245.

Gil, R. and C. A. Ruzzier (2018). The Impact of Competition on "Make-or-Buy" Decisions: Evidence from the Spanish Local TV Industry. *Management Science 64*, 1121–1135.

Grossman, S. J. and O. D. Hart (1986). The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy 94*, 691–719.

Hart, O. and J. Moore (1990). Property Rights and the Nature of the Firm. *Journal of Political Economy 98*, 1119–1158.

Igami, M. (2017). Estimating the Innovator's Dilemma: Structural Analysis of Creative Destruction in the Hard Disk Drive Industry, 1981-1998. *Journal of Political Economy 125*, 798–847.

Krieger, J., D. Li, and D. Papanikolaou (2017). Developing New Drugs. *mimeo*.

Lafontaine, F. and M. Slade (2007). Vertical Integration and Firm Boundaries: The Evidence. *Journal of Economic Literature 45*, 629–685.

Lakdawalla, D. N. (2018). Economics of the Pharmaceutical Industry. *Journal of Economic Literature 56*, 397–449.

Lerner, J. and R. P. Merges (1998). The Control of Technology Alliances: An Empirical Analysis of the Biotechnology Industry. *The Journal of Industrial Economics 46*, 125–156.

Levy, D. T. (1985). The Transactions Cost Approach to Vertical Integration: An Empirical Examination. *The Review of Economics and Statistics 67*, 438–445.

Moorthy, K. S. and I. P. L. Png (1992). Market Segmentation, Cannibalization, and the Timing of Product Introductions. *Management Science 38*, 345–359.

Nahm, J. (2004). Durable-Goods Monopoly with Endogenous Innovation. *Journal of Economics & Management Strategy 13*, 303–319.

Nicholson, S., P. M. Danzon, and J. McCullough (2005). Biotech-Pharmaceutical Alliances as a Signal of Asset and Firm Quality. *Journal of Business 78*, 1433–1464.

Tucker, I. B. and R. P. Wilder (1977). Trends in Vertical Integration in the U.S. Manufacturing Sector. *The Journal of Industrial Economics 26*, 81–94.

Waldman, M. (1993). A New Perspective on Planned Obsolescence. *Quarterly Journal of Economics 108*, 273–283.

Waldman, M. (1996). Planned Obsolescence and The R&D Decision. *RAND Journal of Economics 27*, 583–595.

Williams, H. L. (2013). Intellectual Property Rights and Innovation: Evidence from the Human Genome. *Journal of Political Economy 121*, 1–27.

Descriptive Statistics on the Control Variables (1989-2004)

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Pre-clinical	109,115	0.792	0.406	0	1
Phase I	109,115	0.061	0.239	0	1
Phase II	109,115	0.082	0.275	0	1
Phase III	109,115	0.044	0.204	0	1
Launched	109,115	0.021	0.145	0	1
Experience	109,115	191.436	304.518	1	1,974
PDM	109,115	26.637	12.313	0	42
Scope	109,115	0.097	0.119	0.011	1
TDM	109,115	51.376	24.818	2	81
MSP (with imputations)	109,115	3.952	8.311	0	85.326
MSP (without imputations)	51,439	8.383	10.457	0	85.326

Back

Summary of Drug Development Phases

Development Stage	Description (according to the FDA)
Pre-clinical Trial	Submission of an investigational new drug application for FDA review. Companies need to show pre-clinical testing results on laboratory animals and propose plans for human testing.
Phase I Trial	Usually conducted in healthy volunteers to determine the most frequent side effects, and how the drug is metabolized and excreted. Number of subjects ranges from 20 to 80. Emphasis is on safety.
Phase II Trial	Obtain preliminary data on whether the drug treats a certain disease or condition. Number of subjects ranges from a few dozen to about 300. Continues to evaluate safety and short-term side effects.
Phase III Trial	The FDA and the sponsors meet to determine how large-scale studies in Phase III should be done. Gather more information on safety and effectiveness. Studies different populations, dosages, and combined usage of other drugs. Number of subjects ranges from several hundred to about 3,000 people.

Back

Therapeutic Classification: Panel vs. Compound-Level

Description	Panel Data		Compound-Level	
	Freq.	Percent	Freq.	Percent
A Alimentary tract and metabolism	6,229	5.71	700	6.09
B Blood and blood forming organs	4,167	3.82	410	3.57
C Cardiovascular system	10,924	10.01	975	8.48
D Dermatologicals	2,400	2.20	230	2.00
F Formulations	1,756	1.61	171	1.49
G Genito urinary system and sex hormones	3,460	3.17	358	3.11
H Systemic hormonal preparations (excl. sex hormones and insulins)	2,280	2.09	200	1.74
J Antiinfectives for systemic use	17,476	16.02	1,820	15.84
L Antineoplastic and immunomodulating agents	27,167	24.90	3,084	26.83
M Musculo-skeletal system	5,916	5.42	646	5.62
N Nervous system	19,482	17.85	2,147	18.68
P Antiparasitic products, insecticides and repellents	451	0.41	43	0.37
R Respiratory system	4,662	4.27	460	4.00
S Sensory organs	988	0.91	92	0.80
V Various	1,757	1.61	157	1.37
Total	109,115	100.00	11,493	100.00

Notes: This table reports the main therapeutic class distribution in our sample. The therapeutic categorization used is the main level of the standard Anatomical Therapeutic Chemical (ATC) Classification System developed by the World Health Organization. For a very small share of observations, we cannot map Pharmaproject therapeutic class to ATC, and we used the F group for this Pharmaproject-only group.

Back

Firm Types: Overall, In-house, and Outsource

Originator Firm Type	Overall		In-house		Outsource	
	Freq.	Percent	Freq.	Percent	Freq.	Percent
Biopharmaceuticals	102,986	94.38	81,162	94.78	21,824	92.94
Chemicals	3,406	3.12	2,537	2.96	869	3.70
Health (broad)	2,507	2.30	1,825	2.13	682	2.90
Academia/research/NPOs	136	0.12	78	0.09	58	0.25
CRO/CDMO	80	0.07	31	0.04	49	0.21
Total	109,115	100.00	85,633	100.00	23,482	100.00

Licensee Firm Type	Overall			Outsource		
$1\{\}$ indicators	Obs	Mean	SD	Obs	Mean	SD
Biopharmaceuticals	109,115	0.216	0.412	23,482	0.784	0.412
Chemicals	109,115	0.015	0.123	23,482	0.053	0.223
Health (broad)	109,115	0.012	0.109	23,482	0.038	0.191
Academia/research/NPOs	109,115	0.005	0.069	23,482	0.021	0.143
CRO/CDMO	109,115	0.002	0.046	23,482	0.008	0.089

Back

In-house Dev: Market Share (no imputation) and Patents

	(1) In-house	(2) In-house	(3) In-house	(4) In-house	(5) In-house
Panel B					
Current MSP	0.0126**	0.0107**	0.0107**	0.00949*	0.00971*
	(0.00526)	(0.00522)	(0.00531)	(0.00523)	(0.00528)
EOP1		0.627***			
		(0.195)			
EOP2			0.0135**		
			(0.00555)		
LOP1				0.0435***	
				(0.0118)	
LOP2					0.00129***
					(0.000375)
Observations	51,439	51,439	51,439	51,439	51,439

Back

In-house Dev: Future Market Share (no imputation) & Patents

	(1) In-house	(2) In-house	(3) In-house	(4) In-house	(5) In-house
Panel B					
Future MSP	0.0165**	0.0130*	0.0145*	0.0126*	0.0135*
	(0.00758)	(0.00748)	(0.00779)	(0.00750)	(0.00778)
EOP1		0.976***			
		(0.244)			
EOP2			0.0295**		
			(0.0117)		
LOP1				0.0589***	
				(0.0138)	
LOP2					0.00193***
					(0.000731)
Observations	8,165	8,165	8,165	8,165	8,165

Back

Alternative In-house Defn: Existence & Length of Patents

	(1) In-house	(2) In-house	(3) In-house	(4) In-house
EOP1	0.500***			
	(0.0605)			
EOP2		0.0585***		
		(0.00539)		
LOP1			0.0266***	
			(0.00358)	
LOP2				0.00403***
				(0.000341)
Observations	109,115	109,115	109,115	109,115

Dependent variable is one if a compound is developed in-house by the end of phase I, and zero otherwise. In contrast to our main dependent variable, which indicates whether a compound is developed in-house by the end of phase II, this alternative in-house measure aims to address the concern that the design and nature of a drug may be fixed as early as the completion of phase I testing.

Back

Patent Profile Defined on Finer Therapeutic Classifications

	(1) In-house	(2) In-house	(3) In-house	(4) In-house
EOP1	0.644***			
	(0.0607)			
EOP2		0.0965***		
		(0.0120)		
LOP1			0.0300***	
			(0.00370)	
LOP2				0.00622***
				(0.000766)
Observations	107,098	107,098	107,098	107,098

This table reports the results of constructing drug profile variables based on the second Anatomical Therapeutic Chemical (ATC) level, capturing finer therapeutic classifications. The sample is slightly smaller than in the main analysis, as some observations are dropped due to the more demanding fixed effects of the finer categories.

Back

Robustness Checks: with Firm Type Fixed Effects

	(1) In-house	(2) In-house	(3) In-house	(4) In-house
EOP1	0.503***			
	(0.0603)			
EOP2		0.0599***		
		(0.00537)		
LOP1			0.0271***	
			(0.00358)	
LOP2				0.00411***
				(0.000338)
Observations	109,115	109,115	109,115	109,115

This table reports the results after adding firm type fixed effects to account for potential differences in originator's business models. Categorized firm types include biopharmaceuticals, chemicals, health (broad), academia/research/NPOs, CRO/CDMO.

▸ Firm type summary table ◂ Back

Robustness Checks: excluding CROs/CDMOs

	(1) In-house	(2) In-house	(3) In-house	(4) In-house
EOP1	0.504***			
	(0.0605)			
EOP2		0.0602***		
		(0.00538)		
LOP1			0.0269***	
			(0.00360)	
LOP2				0.00410***
				(0.000340)
Observations	108,846	108,846	108,846	108,846

Back

Robustness Checks: Biopharmaceutical Firms Only

	(1) In-house	(2) In-house	(3) In-house	(4) In-house
EOP1	0.523***			
	(0.0624)			
EOP2		0.0593***		
		(0.00537)		
LOP1			0.0287***	
			(0.00368)	
LOP2				0.00411***
				(0.000339)
Observations	102,986	102,986	102,986	102,986

Back

Robustness Checks: Firm Portfolios

	(1) In-house	(2) In-house	(3) In-house	(4) In-house
EOP1	0.406***			
	(0.0626)			
EOP2		0.0773***		
		(0.0100)		
LOP1			0.0198***	
			(0.00378)	
LOP2				0.00514***
				(0.000568)
Observations	109,115	109,115	109,115	109,115

This table reports results when controlling for additional portfolio measures, including the total number of products in a firm's pipeline in a given therapeutic category each year and the total number of competing products each year in the same therapeutic category of the compound of interest by other firms.

Back

Robustness Checks: Firm Fixed Effects

	(1) In-house	(2) In-house	(3) In-house	(4) In-house
EOP1	0.0396			
	(0.0859)			
EOP2		0.0206***		
		(0.00598)		
LOP1			-0.00275	
			(0.00496)	
LOP2				0.00157***
				(0.000371)
Observations	105,224	105,224	105,224	105,224

This table reports results when including firm fixed effects, and with reduced observations due to projects owned by firms with single/small projects.

Back

Robustness Checks: Development Speed Test

	(1) pre-III	(2) pre-II	(3) pre-I	(4) II-III
In-house	-0.848***	-0.783***	-0.655***	-0.380**
	(0.199)	(0.128)	(0.114)	(0.161)
Observations	593	1,202	1,318	436

This table reports regression in drug-level data using years between phases as the dependent variable and in-house development status as the main covariate of interest. Columns 1-4 have outcome variables as years between pre-clinical to phase III, pre-clinical to phase II, pre-clinical to phase I, and phase II-III, respectively.

Back

Subsample Analysis: by the End of Phase II

	(1) In-house	(2) In-house	(3) In-house	(4) In-house
EOP1	0.526***			
	(0.0607)			
EOP2		0.0612***		
		(0.00559)		
LOP1			0.0277***	
			(0.00359)	
LOP2				0.00425***
				(0.000355)
Observations	102,037	102,037	102,037	102,037

Back

Subsample Analysis: with No Ownership Changes

	(1) In-house	(2) In-house	(3) In-house	(4) In-house
EOP1	0.496***			
	(0.0631)			
EOP2		0.0692***		
		(0.00635)		
LOP1			0.0278***	
			(0.00373)	
LOP2				0.00492***
				(0.000408)
Observations	101,352	101,352	101,352	101,352

Back

Subsample Analysis: Adjust for Potential Censoring

	(1) In-house	(2) In-house	(3) In-house	(4) In-house
EOP1	0.572***			
	(0.0658)			
EOP2		0.0584***		
		(0.00571)		
LOP1			0.0296***	
			(0.00390)	
LOP2				0.00393***
				(0.000348)
Observations	98,842	98,842	98,842	98,842

Back