

NBER WORKING PAPER SERIES

IF YOU HAD ONE SHOT:
SCALE AND HERDING IN INNOVATION EXPERIMENTS

Ashish Arora
Sharique Hasan
William D. Miles

Working Paper 33682
<http://www.nber.org/papers/w33682>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2025

Author names are listed in alphabetical order. There are no relevant financial relationships to report. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

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 They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2025 by Ashish Arora, Sharique Hasan, and William D. Miles. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

If You Had One Shot: Scale and Herding in Innovation Experiments

Ashish Arora, Sharique Hasan, and William D. Miles

NBER Working Paper No. 33682

April 2025

JEL No. L1, L2, O31, O32, O33

ABSTRACT

Solving complex problems- in medicine, engineering, and other technological domains- often requires exploring multiple approaches, particularly when significant uncertainty exists about which one will lead to success. Conventional wisdom assumes that having many experimenters independently decide which approaches to pursue increases diversity and, thus, also the chances of finding a solution. However, if experimenters herd toward the most promising approach, this convergence may reduce diversity and thus the likelihood of solving the problem. In this paper, we develop a simple model to show that, holding the total number of experiments constant, markets dominated by a few large-scale experimenters-firms conducting multiple experiments- explore more diverse approaches than markets with many single-shot experimenters. Single-shot experimenters tend to converge on the most promising approach, while multi-experimenters are more likely to diversify to avoid the correlation inherent in pursuing multiple experiments within the same approach. We test our model's predictions using data from pharmaceutical R&D. Our analysis shows that increasing the average number of experiments per firm by one unit raises target diversity by over three standard deviations. In turn, a one-standard deviation increase from the mean in target diversity boosts the likelihood of at least one experiment reaching Phase 1 clinical trials by 25.9 percentage points. Our findings inform policies for the optimal allocation of experiments across firms to maximize approach diversity and market-level success.

Ashish Arora
Fuqua School of Business
Duke University
Box 90120
Durham, NC 27708-0120
and NBER
ashish.arora@duke.edu

William D. Miles
100 Fuqua Drive
Durham, NC 27708
United States
will.miles@duke.edu

Sharique Hasan
Duke University
Fuqua School of Business
Box 90120
Durham, NC 27708-0120
sh424@duke.edu

A data appendix is available at <http://www.nber.org/data-appendix/w33682>

Replication materials are available at <https://github.com/willdmiles/if-you-had-one-shot>

1 Introduction

From advancing healthcare to addressing climate change, solving society’s most pressing challenges requires innovation and extensive experimentation. Breakthroughs often emerge when a wide range of different approaches are explored (Gross and Sampat, 2023). In the context of technological progress, an approach can be understood as a structured hypothesis or method designed to explore specific pathways, mechanisms, or combinations of variables in solving a problem (Sorenson, 2024). For instance, the development of statins was made possible by years of experimentation, including Akira Endo’s groundbreaking work on fermented natural products to reduce cholesterol. Similarly, the invention of the transistor arose from systematic, extensive testing of material and design combinations across multiple labs, culminating in the 1947 breakthrough by John Bardeen, Walter Brattain, and William Shockley. Thomas Edison’s quest for the ideal filament for the light bulb involved testing thousands of combinations of materials and configurations. These examples illustrate how diverse experimentation increases the odds of solving complex problems.

Conventional wisdom holds that increasing the number of experiments leads to greater variety in the approaches used to solve complex problems (Prager, 2021; Hong and Page, 2004). Underpinning this conventional wisdom is the belief that involving many small-scale experimenters is a more effective way to achieve approach diversity than relying on a few larger experimenters conducting multiple experiments each (Cohen and Klepper, 1992). The logic is simple: more solvers are assumed to independently explore a wider range of approaches by bringing in diverse perspectives (Thomke et al., 1998). However, understanding the incentives of firms is critical to determining whether they will explore new approaches or herd toward the most promising one. What should a firm do if it had *only* one shot? All else equal, a firm in this position is likely to focus on the most promising approach, as this maximizes its perceived chance of success (Bikhchandani et al., 1992). In the aggregate, however, the concentration of efforts on the same approach diminishes diversity across experiments, ultimately reducing the overall likelihood of solving the problem. This raises the question: under what conditions can a greater diversity of approaches be achieved, and how does this

influence the likelihood of success?

In this paper, we develop a simple model to examine how firms decide which approaches to pursue, distinguishing between multi-experiment firms and one-shot experimenters. In our model, experiments face two sources of failure: an approach may be a dead end, or its implementation may fail even if the approach is otherwise viable. Multi-experiment firms reduce the correlation between their outcomes by spreading experiments across separate approaches (with independent likelihoods of being viable), increasing the likelihood of at least one success. In contrast, one-shot experimenters tend to converge on the most promising approach, independently herding into the same strategy and reducing diversity at the market level. As a result, our model predicts that markets with a few multi-experiment firms will generate greater diversity in approaches than otherwise similar markets dominated by many one-shot experimenters. Diversity results in a higher probability of the market finding at least one successful solution, though the average experiment is less likely to succeed: Diversity reduces the probability of complete failure in return for a lower probability of success per experiment. This tradeoff between average and aggregate success is central to our argument and has important implications for both firm strategy and innovation policy.

To test the implications of our model, we use a comprehensive dataset on early-stage pharmaceutical drug development from 1996 to 2023, which includes 49,866 drug development projects led by 3,845 firms across 241 therapeutic classes. These data capture the preclinical testing of drug-target combinations, where a therapeutic class-target pair represents an approach, and a firm’s trial of that pair constitutes an implementation attempt. This framework allows us to empirically examine two key hypotheses: (1) within a therapeutic class, as the average number of experiments conducted per firm increases, the diversity of approaches—measured as the variety of targets drugged—will rise; and (2) greater diversity of approaches will decrease the average success rate per trial while increasing the likelihood of at least one successful outcome within a therapeutic class.

Our empirical analysis supports these hypotheses. Markets with a one-unit increase in average experimenter scale exhibit a 3.4 standard deviation increase in target diversity. Additionally, a higher level of diversity is associated with a trade-off: a 20.3 percentage

point increase in the likelihood of at least one experiment progressing to Phase 1 clinical trials, alongside a decline in individual experiment success rates. These results show that the experimenter scale can foster approach diversity and increase the probability of success in solving complex problems with uncertain solutions.

While the evidence supports our main theoretical predictions, other explanations could also account for the patterns we observe. For example, markets may differ in size, technological opportunity, or firm capabilities, which may affect the average scale of experiments and the likelihood of success. However, none of these can explain the core empirical associations we document, namely that the average experimenter scale is positively associated with diversity and positively associated with the probability of at least one successful solution but *negatively* associated with average success per experiment.

We control for the total number of targets, the number of firms, and the total number of projects in a therapeutic area (a higher level of aggregation than a market), along with therapeutic area and time fixed-effects. Our results are also similar when we additionally control for the rate of target discoveries within disease classes. While we find that single-experiment firms do indeed introduce a significant proportion of new targets, this does not diminish the relationship between experimenter scale and diversity. Furthermore, contrary to the assumption that one-off experiments are exclusive to startups, both large public firms and startups actively engage in single experiments. Finally, to account for differences in firm capabilities and experience, we include the average age of firms in a therapeutic class as a control variable and find consistent results (Krieger et al., 2022).

Taken together, these additional analyses support our main findings: experimenter scale plays a crucial role in enhancing approach diversity and reducing the likelihood of outright failure at the market level. Nevertheless, small-scale experimenters play a crucial role in introducing new targets to experimentation, with firms conducting only one experiment being more likely to explore novel targets than multi-experimenters. These complementary findings highlight a critical dynamic: while markets with greater average experimenter scale foster diversity by leveraging the broader experimentation efforts of large-scale firms, they also depend on the exploratory efforts of small-scale experimenters to expand the frontier of

novel targets.

This paper makes three key contributions. First, to the literature on experimentation, we show that the scale of experimentation—how many shots each firm takes—fundamentally shapes market-level diversity and success. This shifts focus from the effectiveness of individual experiments to the collective architecture of experimentation. While the literature has primarily focused on the effectiveness of experimentation (Koning et al., 2022), our paper shifts the focus to the scale of experimentation and explicitly relates it to market-level outcomes. We thus link experimentation to technology policy, particularly regarding the optimal societal organization of collective experimentation (Nelson, 1961).

Second, we offer a conceptual contribution by distinguishing between two components of an experiment: the approach and the implementation attempt. Camuffo et al. (2024) examine how scientific training can help entrepreneurs design better experiments. Implicitly, much of the literature on experimentation, including studies by Koning et al. (2022) and Gans et al. (2019), do not distinguish between the implementation and approach, instead bundling them into a single concept. We argue that these are distinct (Bryan and Lemus, 2017; Dasgupta and Maskin, 1987). An approach refers to a hypothesis or theory about the cause of a problem, suggesting potential ways to solve it (Sorenson, 2024). In contrast, an implementation is a specific attempt to solve the problem based on the hypothesized cause. For success, both the approach must be viable and the implementation must be effective. As we demonstrate, the shared risk of failure across all implementation attempts using the same approach can lead to excessive herding.¹

Third, our empirical results emphasize the distinct ways in which large- and small-scale experimenters affect diversity. Small-scale experimenters contribute by introducing novel approaches, thereby expanding the portfolio of available approaches. Large-scale experimenters, on the other hand, add diversity by more thoroughly investigating known but understudied approaches. This finding ties to the literature on technological trajectories, suggesting that

¹Throughout the paper, we distinguish “approach” (a hypothesis about a solution pathway, such as targeting a specific protein) from “implementation” (a particular experimental effort based on that approach, such as testing a specific compound). This distinction is essential to our model, as it helps explain how correlated failures can arise even from independent trials.

markets with small-scale experimenters may underexplore some trajectories—herding towards the ones that are most likely to succeed (Ciarli et al., 2021; Nelson et al., 2023; Tan, 2023).

2 Related Literature

In this section, we briefly review the literature relevant to our paper, connecting ideas from three domains: experimentation and its market-level implications, firm-level incentives to innovate, and the relationship between firm size and the diversity of approaches. These literatures provide a foundation for understanding how firms navigate their choice of experimental strategy and how their choices influence market-level outcomes. These literatures motivate our model in Section 4, which explores how firm-level experimentation strategies shape market-level outcomes.

The first literature we draw on is the recent work on experimentation, which has thus far focused on the design, effectiveness, and decision-theoretic implications of experimental strategies. For example, Koning et al. (2022) and Gans et al. (2019) emphasize how decision-makers design experiments to gather information that improves subsequent investments, product designs, or market strategies. In these studies, experimentation is valued primarily for reducing uncertainty and enabling better decision-making. However, our work shifts the focus to settings where experiments yield direct private payoffs but have broader market-level consequences. For instance, in our empirical context of clinical trials, additional successes do not increase the overall benefit once one firm succeeds, as the total value is shared among all successful firms. This market feature creates a gap between what benefits individual firms and what is best for society. Moreover, this gap depends on the composition of the experimenters. Thus, we extend the literature to explore market-level technological diversity and innovation.

Our paper also builds on the theoretical literature examining how firms choose experimental strategies. Our model shares similarities with Nelson (1961). However, while Nelson assumes a single attempt per approach, we do not impose such constraints, allowing for a

more flexible representation of experimentation under uncertainty. Bryan and Lemus (2017) also explore firms' choices among competing innovation approaches. Unlike them, we examine the case where some firms may follow multiple approaches but approaches could be dead ends. This extension captures the inherent uncertainty in experimental settings, where outcomes within a given approach correlate. We find that the market may have too little diversity, which aligns with Dasgupta and Maskin (1987), who demonstrate when firms ignore the effect of their investments on others, their research portfolios will be too similar to each other.

The second domain we connect is the literature on firm-level incentives to innovate, particularly how private incentives can lead to herding and clustering in specific technological trajectories (Lieberman and Asaba, 2006; Krieger, 2021). Herding has been extensively studied, notably in the work of Bikhchandani et al. (1992), which examines how agents may disregard private information to converge on specific actions based on observed behavior. In contrast, our framework focuses on static beliefs, where herding arises not from belief updating but from private incentives to pursue the most promising approaches. This behavior reflects static, payoff-driven incentives to converge on the most promising approach—distinct from informational cascades driven by belief updating. These dynamics highlight how private incentives can contribute to reduced diversity and suboptimal outcomes at the market level, underscoring the tension between individual and collective rationality in innovation.

Finally, our work speaks to the relationship between firm size and the diversity of approaches. Larger firms, which can conduct multiple experiments simultaneously, face fundamentally different trade-offs than smaller firms constrained to one-off experiments. Cohen and Klepper (1992) argue that concentrating experiments within a few firms reduces the diversity of approaches. Our model diverges from this stream of research by considering non-additive payoffs, where the incremental value of success diminishes significantly with each additional success. This non-additivity affects diversity and innovation at the market level.

3 Experiments and Approaches

3.1 What is an approach?

For any technical problem where a solution is not yet known, there may be multiple potential approaches to solve it. Each approach represents a theory about the underlying causal structure that explains outcomes as a function of a set of inputs (Sorenson, 2024). Consider the challenge of building a quantum computer. A key decision is how to create quantum particles, known as qubits. Different firms have adopted fundamentally distinct approaches to this problem. For example, IonQ creates physical qubits using trapped-ion technology.² In contrast, Microsoft focuses on topological qubits, where quantum information is stored in a collective group of particles rather than in the state of a single one.³ Meanwhile, Google employs superconducting circuits as synthetic qubits, encoding and manipulating quantum information in superconducting circuits cooled to extremely low temperatures.⁴ The choice of approach represents a significant strategic bet. That is, among innovative companies working to solve a given problem—whether it be curing a disease, developing generative AI algorithms, or developing self-driving cars—the choice of approach is one of the most critical decisions they make.

3.2 Implementing experiments within an approach

The second component of experimentation is implementation, which involves putting an approach into action through specific experiments or trials. While firms may adopt the same approach, their success often hinges on how effectively they implement it. For example, in the case of quantum computing, Google is not the only company using superconducting qubits. Other firms, such as IBM, Rigetti, and Intel, have also chosen this approach but differ in their implementations of superconducting circuits. Even if superconducting circuits prove to be a viable pathway for building quantum computers, not all firms pursuing this

²For more details, see <https://ionq.com/company>

³For more details, see <https://azure.microsoft.com/en-us/solutions/quantum-computing/technology/>

⁴For more details, see <https://quantumai.google/quantumcomputer>

approach will succeed. Their success will ultimately depend on the effectiveness of their specific implementations.

3.3 Two sources of experimental uncertainty

The distinction between approaches and implementations highlights two sources of uncertainty in any experiment, both of which are critical to understanding the outcomes of innovative efforts. The first is *approach uncertainty*—whether the underlying hypothesis behind the approach is valid and capable of leading to an effective solution. The second is *implementation uncertainty*—whether the firm can successfully execute the necessary steps, such as designing a superconducting circuit or synthesizing a lead compound to drug a target. Figure 1 maps experimental outcomes along two dimensions: whether the approach is correct or wrong, and whether the implementation is successful or unsuccessful.

[Figure 1 about here.]

The upper-left quadrant represents the ideal scenario, where the approach is valid and the specific implementation is successful. Examples of such experiments include the use of CRISPR technology to edit genes. In the top-right quadrant, the approach is valid, but a specific implementation fails. An example of this is the Wright brothers' early attempts to invent the airplane. While their concept for heavier-than-air flight was correct, several implementations were unsuccessful. The bottom row highlights cases where the approach is incorrect in the first instance, i.e., a dead end. The bottom-left quadrant is an interesting case where, despite a dead-end approach, a successful solution is found, such as the serendipitous discovery of penicillin by Alexander Fleming. Another example is the radio wave detector, the “audion”, invented by Lee DeForest, which, though it worked, relied on the incorrect belief that a low-pressure gas inside the glass bulb was required.⁵ Lastly, in the lower-right quadrant, the approach is wrong, and the subsequent implementation efforts also fail. For instance, early tungsten filament lightbulbs were plagued by blackening on the inside of the bulb, and the proposed approach was to improve the vacuum inside the bulb. Implementation efforts

⁵<https://en.wikipedia.org/wiki/Audion>

failed, and it later turned out that the approach was also incorrect before Irving Langmuir suggested filling the bulb with an inert gas. He correctly hypothesized that oxidation was not to blame. Instead, the heated filament itself was emitting electrons that were deposited on the glass surface, blackening it. Inert gases inside the bulb scattered the electrons, solving the problem.

3.4 Case Study: Experimentation in Alzheimer’s Disease

The distinction between approaches and implementation underscores an important implication: outcomes of experiments relying on the same approach are correlated. This connection highlights a critical challenge for innovation when many firms converge on a single hypothesis or theory. A prominent example is drug development in Alzheimer’s Disease (AD). While numerous hypotheses exist about the causes of AD and potential ways to prevent its onset, an effective cure remains elusive. Compounding this difficulty is the considerable crowding around only a few approaches. Table 1 reports the number of AD drug development projects initiated by leading hypotheses between 1998 and 2008. Nearly 40% of projects focused on the beta-amyloid hypothesis, and almost 60% targeted just two hypotheses: beta-amyloid and cholinergic.

Although there have been three FDA-approved drugs to target beta-amyloids, their benefits are very modest.⁶ These drugs are known to cause serious side effects, such as brain bleeds, and have minimal impact on slowing the progression of AD in very early-onset patients.⁷ Adding to the concern, recent research has raised doubts about the legitimacy of key beta-amyloid science (Piller, 2024, 2022). If the beta-amyloid hypothesis is ultimately incorrect, the effectiveness of a firm’s implementation becomes irrelevant: *all drug candidates relying on beta-amyloid will likely fail.*

[Table 1 about here.]

⁶One of these drugs—Aducanumab—was discontinued in late 2024 after its creator Biogen “faced scrutiny of its pursuit of approval and the steep list price it set for a drug that many doctors and researchers said wasn’t fully proven to work.” <https://www.wsj.com/tech/biotech/biogen-ends-aduhelm-program-in-shift-of-alzheimers-resources-163c897a>

⁷<https://www.wsj.com/articles/new-alzheimers-drug-shows-positive-results-but-side-effects-11669766449>

Which firms are crowding into the beta-amyloid and cholinergic hypotheses? Figure 2 examines projects initiated between 2007 and 2008, comparing Pfizer—which launched six distinct AD drug development projects during this period—to the 18 firms that each initiated only one project in the same timeframe. This figure illustrates how a large-scale experimenter can contribute to greater diversity in approaches. Notably, only 2 out of Pfizer’s 6 projects (33%) focused on beta-amyloid drugs. In contrast, among the single-experiment firms, 8 out of 18 projects (44%) targeted beta-amyloids. These results suggest that larger experimenters, such as Pfizer, are more likely to diversify their approaches, reducing the risk of correlated failures, whereas smaller experimenters—who make their decisions independently—tend to crowd into popular hypotheses, increasing systemic vulnerability to approach-level failure.

[Figure 2 about here.]

4 Model

We begin by formalizing the core elements of our setup, then analyze how firm-level experimental scope affects approach diversity and the likelihood of successful innovation at the market level.

4.1 A model of market structure and approach diversity

In this section, we develop a model examining how firms choose between experimental approaches to solve innovation problems with uncertain solutions where a single success is sufficient for the market. The model explores how market structure—the presence of many small versus few large experimenters—affects the diversity of approaches pursued and market-level outcomes. Our model has four key elements that capture the essential features of collective experimentation in innovation. First, we distinguish between approaches and implementations. Second, even the most promising approach is not certain to be viable, i.e., there is a non-zero probability that it could be a dead-end. Third, we assume that the combined payoff to all successful experiments is fixed and divided between all successful firms. In particular,

a successful experiment yields a positive private payoff, but diminishes in proportion to the number of successful firms, and a second success for the same firm has no incremental value. Finally, beliefs are static and there is no learning.

4.2 Approaches versus implementations

Suppose there are two potential approaches $\{a, b\}$ from which a firm can choose to solve a problem. Let π_a and π_b represent the known probabilities that approach a and b , respectively, are viable solutions to the problem. We represent the probability that the implementation of approaches $\{a, b\}$ is effective with $\{p_a, p_b\}$. That is, an experiment in approach a has a probability of success of $p_a\pi_a$, and the probability of success with approach b is $p_b\pi_b$. There is a common belief that a is the more promising approach so that $\pi_ap_a > \pi_bp_b$. Suppose a firm conducts two experiments, one in each approach. Let $x_a = 1$ if the experiment in a succeeds and $x_b = 1$ if the experiment in b succeeds. Since the approaches are independent (the viability of approach a is not related to the viability of approach b), we have that $Cov(x_a, x_b) = \pi_a\pi_bab - (\pi_aa)(\pi_bb) = 0$. Consider a firm that conducts two experiments, both using approach a . Because the experiments use the same approach, their expected outcomes are correlated because $Cov(x_a, x_b) = \pi_aa^2 - (\pi_aa)(\pi_aa) = a^2\pi_a(1 - \pi_a) > 0$.

Let v be the value from a successful experiment and c be the cost. All firms derive the same value from an experiment and face the same cost.⁸ As noted earlier, only one success matters and the incremental benefit from the second success is zero (Dasgupta and Maskin, 1987). Thus, for a firm with two experiments, two successes yield the same payoff as one success. If two independent firms succeed, they split the payoff. A firm with two experiments has an incentive to hedge by choosing a different, even if less promising, approach for its second experiment.

⁸Allowing some firms to appropriate higher rents from the same innovation (Cohen and Klepper, 1996a) affects the number of experiments (i.e., the decision to enter) but not the choice of approach.

4.3 Single and Multi-Experiment Firms and Market-Level Approach Diversity

If firm i uses approach a , its expected payoff is $vp_a\pi_a - c$, and if it uses approach b , its expected payoff is $vp_b\pi_b - c$. Given our assumptions, the firm should choose approach a for their experiment if $vp_a\pi_a - c \geq 0$, and otherwise not enter at all. We now analyze the choice of approach for the next experiment under two scenarios: first, where a separate firm j considers the second experiment, and second, where firm i itself conducts the second experiment.

Case 1 – Two single-experiment firms: A separate firm j observes firm i 's decision to use approach a , but firm i 's experiment is ongoing, so outcomes are unknown. Given this information, and that only one experiment can succeed and earn rents, firm j 's expected payoff from also using approach a is $v\pi_a \left[p_a(1 - p_a) + \frac{1}{2}p_a^2 \right] - c$. Intuitively, firm j 's payoff is the probability that only its experiment is successful, plus half the probability that both firms are successful (since there is an even chance that firm j wins when both succeed). Note that both probabilities depend on π_a , the probability that approach a is viable. If firm j instead decides to use approach b , its expected payoff is $v \left[\pi_b p_b(1 - \pi_a p_a) + \frac{1}{2}\pi_a \pi_b p_a p_b \right] - c$. The difference between these expected payoffs, by $\Delta_s = v \left[\pi_a p_a \left(1 - \frac{1}{2}p_a \right) - \pi_b p_b \left(1 - \frac{1}{2}\pi_a p_a \right) \right]$, where Δ_s represents the single-experiment firm's incentive to 'herd' into approach a , following the crowd despite its being the more crowded path.

Case 2 – One multi-experiment firm: Instead of firm j , suppose firm i is considering a second experiment. As before, we assume the outcome of the first experiment is unknown. If firm i chooses to conduct its second experiment in approach a also, its expected payoff becomes $v\pi_a (1 - (1 - p_a)^2) - 2c$. Since firm i already has an experiment in approach a and there is no additional benefit to a second success, this payoff represents one minus the probability of failing in both experiments, minus the cost of two experiments. If firm i instead conducts its second experiment in approach b , its expected payoff is $v \left[\pi_a p_a(1 - \pi_b p_b) + \pi_b p_b(1 - \pi_a p_a) + \pi_a \pi_b p_a p_b \right] - 2c$. Taking the difference between these expected payoffs, we can derive an expression for firm i 's incentive to use approach a in its

second experiment: $\Delta_M = v [\pi_a p_a (1 - p_a) - \pi_b p_b (1 - \pi_a p_a)]$

We can now compare each firm's incentive to crowd into the dominant approach: the multi-experiment firm's incentive Δ_M versus the single-experiment firm's incentive Δ_S .

$$\Delta = \Delta_M - \Delta_S = \frac{1}{2} v \pi_a p_a [\pi_b p_b - p_a] \leq 0 \iff \pi_b p_b \leq p_a \quad (1)$$

A multi-experiment firm is less likely to crowd into the dominant approach than two single-experiment firms when $p_a > \pi_b p_b$, which always holds given our assumption that $\pi_a p_a > \pi_b p_b$ (approach a is more promising). That is, a single firm conducting two experiments is more likely to try different approaches than two separate firms.

This simple two-experiment case demonstrates that approach diversity increases when one firm conducts both experiments rather than when two firms each conduct one. This occurs because *a multi-experiment firm maximizes its chance of at least one success, reducing correlation in outcomes by trading off success probability per attempt.*

Proposition 1 *Holding the number of experiments constant, a firm pursuing multiple experiments will explore a greater variety of approaches compared to an equivalent number of projects, each pursued by a separate firm.*

Given the choices of approach for single and multi-experiment firms, the intuition for the expected outcomes is straightforward. If single-experiment firms are more likely to use the most promising approach in their experiments, each has a probability of success is $\pi_a p_a$, whereas for a multi-experiment firm it is $\frac{1}{2}(\pi_a p_a + \pi_b p_b) < \pi_a p_a$. That is, we expect a higher share of individual experiments to succeed than those of multi-experiment firms.

Proposition 2 *Small-scale experimenters are more likely to be successful in an individual experiment than large-scale experimenters. Therefore, markets with a higher share of small-scale experimenters will have a higher share of successful experiments.*

However, what matters for market outcomes is maximizing the probability of at least one successful experiment. A multi-experiment firm will diversify only when that maximizes this probability, i.e., when $\Delta_M \leq 0$. Yet the expression for the second independent experiment,

Δ_S , is strictly greater than Δ_M . Thus, for some parameter values, we have $\Delta_S > 0 > \Delta_M$. In this range, two single-experiment firms would both use the dominant approach, while one firm conducting both experiments would use different approaches. Critically, the probability of at least one success would be *lower* with two single-experiment firms.⁹ Outside this range, the approach choices would be identical. Put differently, a multi-experiment firm would choose two distinct approaches when this maximizes the probability of at least one success. More formally, the market-level probability of at least one success is maximized with two distinct approaches if and only if:

$$\Delta_M = v [\pi_a p_a (1 - p_a) - \pi_b p_b (1 - \pi_a p_a)] \leq 0 \iff \frac{\pi_b p_b}{\pi_a p_a} > \frac{1 - p_a}{1 - \pi_a p_a} \quad (2)$$

Because $\pi_a p_a > \pi_b p_b$, a *necessary* condition for approach diversity with a multi-experiment firm is that $\pi_a < 1$, as otherwise the inequality in equation 2's second line cannot hold. When $p_a = 1$, the inequality must hold, providing a sufficient condition. This indicates that uncertain approaches combined with effective implementation increase the likelihood that markets with diverse approaches yield a higher probability of at least one success.

Proposition 3 *Markets with greater approach diversity in experimentation have a higher probability of at least one success.*

[Table 2 about here.]

In Appendix B, we generalize to multiple firms and show that the results continue to hold. We also show that when competition between two successful firms dissipates rents, firms will switch to choosing the less promising approach sooner. That is, the difference in incentives to diversify between single and multi-experiment firms will shrink when the total private value of innovation shrinks with the number of successful innovators.

Generalizing beyond two approaches and two possible experiments to multiple approaches and multiple firms involves considerations of competition and order of entry into experimentation, which is not analytically tractable. Simulation results reported in the appendix

⁹As an example, consider $\pi_a = 0.5, \pi_b = 0.4, p_a = 0.5, p_b = 0.5$. The probability of at least one success with two single-experiment firms is 0.375, whereas with a single firm conducting both experiments, it is 0.4.

C indicate that the main intuition is robust to extending to multiple potential entrants and more than two approaches. An increase in the share of multi-experiment firms increases market diversity, lowers average success rates, and lowers the probability of total failure (i.e., increases at least one market-level success).

5 Data & Methods

5.1 Empirical Context: Early-stage drug development

The ideal empirical setting for us is one where there are a variety of possible approaches to a complex, technical problem, and firms enter with one or more experiments. One key requirement is that experimentation must be expensive. If experiments were cheap, all firms would naturally pursue multiple approaches, as in settings like software development ‘a la Koning et al. (2022).

Another desirable feature is that more successes do not create more private value in the aggregate and that competition among successful firms does not fully dissipate the total private value either. Therefore, we want to focus on situations where the incremental value of an additional success is small, both privately and in the aggregate, and where successful firms can differentiate sufficiently to limit rent dissipation through competition.

We use early-stage drug development as the empirical testbed for our model. The development of novel therapeutic drugs is a technologically intensive, high-cost industry where, in many cases, one drug captures a significant majority of the market in a disease area.¹⁰ The potential reward from success is substantial, but so is the cost of experimentation (DiMasi et al., 2016).

The drug development process can take up to 15 years (Hughes et al., 2011) and consists of three main stages: discovery, pre-clinical research, and human clinical trials (Frankel et al., 2023). Data on the discovery process are typically unavailable to archival researchers since these contain secret developments that firms do not want their competitors to know about and will, therefore, not leave a paper trail (Krieger et al., 2022). However, progression to

¹⁰This is true at least during the first 15-20 years that the drug is on the market and still has patent protection.

pre-clinical research is increasingly reported. The choice to move a project into pre-clinical testing is a good approximation for the conditions of our model. Namely, it involves an important choice about what biological target to drug, which is largely informed by scientific knowledge in the public domain (Knowles and Gromo, 2003). Evaluating the promise of a target has also benefited from the recent surge in the availability of big data (Chen and Butte, 2016) and machine learning algorithms (Dara et al., 2022). In the context of our theoretical framework, we define an approach as the choice of therapeutic target (Thomke et al., 1998; Tranchero, 2023), and an implementation as the action of testing the drug-target combination in a preclinical trial. An experiment should be thought of as choosing a target to drug in a pre-clinical trial.

The first data source we use is Pharmaprojects, a commercial database curated by Citeline. Pharmaprojects provides comprehensive global drug research and development history. Drugs are tracked from early pre-clinical development through market launch and are retained regardless of their outcome. New drugs and events for existing drugs are added to the database daily after a thorough editorial process governed by industry experts who scour thousands of sources in the public domain (e.g., press releases).

We collect data from the Pharmaprojects Trends database, which provides annual updates for the status of drug development projects since 1995.¹¹ A status corresponds to a clinical trial phase or the drug being discontinued or launched. It is important to distinguish between a lead compound and a drug development project. For our purposes, a drug development project concerns the clinical trials (development) for a lead compound intended for a particular therapeutic class.¹² We highlight this distinction as it is common for a lead compound to be in clinical trials for more than one therapeutic application. The raw data contain 46,390 distinct lead drug compounds, which correspond to 98,784 drug development projects (i.e., a lead compound is, on average, in development for two therapeutic indications). We keep projects where (i) we can identify the year in which pre-clinical trials began,¹³ (ii) the

¹¹These data were downloaded in August 2024.

¹²In Pharmaprojects, a therapeutic class most closely resembles the Anatomical Therapeutic Chemical (ATC) Classification Level 3, which represents chemical, pharmacological, or therapeutic subgroups.

¹³Furthermore, although the data begin in 1995, we do not have the information to deduce whether projects that first appear in 1995 also began in that year. We thus drop any projects that first appear in the data in 1995. We

organization leading the project is a for-profit firm, and (iii) projects where we can identify the biological target that the drug is intended to act upon. These data also allow us to observe important outcomes for each project, namely if the project advances to phase 1 clinical trial and if the project results in a drug launch. After these cleaning steps, we have 49,866 drug development projects started by 3,845 firms over 28 years between 1996 and 2023 in 241 therapeutic classes.

We also collect data from two additional sources. First, from *Pitchbook* we collect data on firm founding years and ownership status. We use a fuzzy matching algorithm to match Pharmaprojects to Pitchbook by firm name (see Appendix D.2 for an explanation of the match and its accuracy). Second, we collect data from the GWAS catalog, a record of all scientific publications in top-tier journals reporting the discovery of new target-disease correspondences (Tranchoero, 2023). We use these data to measure the rate of new target discovery across therapeutic classes over time.

5.2 Variable Construction

Recall that our model, though based on individual firm-level behavior, addresses market-level rather than firm-level outcomes. Diversity at the individual firm level has little meaning for single-experiment firms. More substantively, the rate of technological advance depends on whether at least one of the experiments launched in the market succeeds.

Our unit of analysis is the market, which we define as a therapeutic class-year(i, t) observation. The choice of time period within which to define a market is somewhat arbitrary. We define market observations within a year, as this will be most consistent with the no updating of beliefs within an observation period: a firm that starts two projects in the same therapeutic class year does not know the outcome of the first project before starting the second. In Appendix E, we replicate our baseline results for two- and 5-year windows and find qualitatively similar trends. We collapse the project-level data to the therapeutic class-year level and then construct market-level variables of interest. We only include therapeutic class-years with (i) at least two firms and (ii) at least two approaches, because when only

also drop projects that began in 2024 as the data are incomplete for this year.

one target is in use, there is, by construction, no diversity in approaches.

Tables 3 and 4 present descriptive statistics and the correlation matrix for the variables in our market-level dataset. The dataset comprises 2,523 observations at the therapeutic class-year (market) level, spanning 139 therapeutic classes. On average, each market features 9.8 distinct approaches deployed across 11.6 experiments, 13.8% of the experiments are initiated by multi-experiment firms, and the average firm conducts 1.2 experiments.

[Table 3 about here.]

[Table 4 about here.]

Market Diversity of Approaches. We measure the market-level diversity of approaches using Shannon entropy (Shannon, 1948), a popular measure used in ecology to quantify the abundance and evenness of species present in a community (Margalef, 1958). Shannon entropy has also recently appeared in studies of the science of science to quantify the certainty of knowledge (Kang et al., 2024). Formally, *Target Diversity* is defined by:

$$Target\ Diversity_{i,t} = - \sum_{k=1}^n p_k \ln p_k$$

Where p_k is the proportion of projects using target k for all drug development projects in disease i started in year t .¹⁴

Success of Experimentation. We create two success measures. We define the success of a drug development project as (i) whether the project advances from pre-clinical experimentation to phase 1 clinical trial and (ii) whether the project results in a drug being launched in the market. We look at these two milestones as representing related but different types of success. Progression to phase 1 clinical trials can be interpreted as a predominant signal of *technical* success and is our preferred outcome for testing our theory. A project will only progress to human trials if the preclinical trial provides substantial evidence of safety and efficacy. While drug launch also signals technical success insofar as the drug has been proven safe and effective in humans, it also represents *commercial* considerations. The cost

¹⁴Using the Herfindahl–Hirschman index gives very similar results, as shown in Appendix Table ??.

of bringing a drug rises steeply through the latter stages of human clinical trials. Hence, a firm’s decision to proceed with later stages also depends on economic factors such as expected competition from drugs by other firms in the market and in their development pipeline.

We create two variables based on these measures: the first is a dummy indicator equal to 1 if in market-year (i, t) *at least one* project succeeds, and the second measure is the *share* of successful projects.

Experimenter Scale. The main independent variable in our analysis captures the distribution of experimenter scale within a market. Our preferred measure, *Average Experimenter Scale*, is defined as the average number of drug development projects initiated (i.e., beginning preclinical trials) by firms in therapeutic class j in year t . We also create the variable *Multi-Experiment Share*, which represents the share of firms in market-year (i, t) that initiate preclinical trials for two or more distinct projects. These measures are highly correlated, with a Pearson correlation coefficient of 0.82. Thus, we do not report results with *Multi-Experiment Share* as they are similar to specifications that use *Average Experimenter Scale*. See also Appendix A for a discussion of measurement error and its implications for our results.

Discovery. The variable $\ln(\textit{Discovery})$ is the natural logarithm of the count of publications in GWAS in year t that are the first to report a relationship between a target and therapeutic class i . The GWAS catalog only reports the disease and/or trait that a particular study addresses, so to match these data to Pharmaprojects, we create a correspondence between diseases/traits and therapeutic classes (the details of this correspondence are explained in Appendix D).

Firm Age. We collect firm founding year data from Pitchbook, which we match by firm name to Pharmaprojects (see Appendix D). We create the market-level measure *Average Firm Age*, which is the average age of all distinct firms that start preclinical experimentation for a new drug development project in therapeutic class i , year t .

Control Variables. In all specifications, we include $ATC-1 \times Year$ fixed effects (Branstetter et al., 2022). ATC-1 represents the first level in the anatomical therapeutic class classification and most closely resembles distinct industry classifications, such as Neurology, Cardiovascular, and Dermatologicals. This interacted fixed effect thus captures time-varying differences

in industry-specific characteristics, such as technology or demand shocks.

The number of approaches and experiments dictates the range of possible values for Shannon entropy. For example, with two approaches and two firms, the only possible value of Shannon entropy is $-(\frac{1}{2} \ln \frac{1}{2} + \frac{1}{2} \ln \frac{1}{2}) = 0.693$. But with ten firms, where each firm does one experiment, Shannon entropy can range from $-\ln \frac{1}{2} = 0.693$ where five firms use one approach and five firms use another, to $\ln 10 = 2.303$ where each firm uses a different approach. Additionally, the number of firms in a market may influence the average size of any experimenter. For example, in a saturated market with low demand but many competitors, firms may be discouraged from pursuing subsequent projects. Accordingly, we include a set of market structure controls, which in all tables are denoted as *Market Structure Controls*. This includes the number of firms, targets, and projects started for each therapeutic class-year.

6 Results

We present our results in four parts. First, we provide correlational evidence for the three propositions derived from our model. We then discuss alternative explanations and how we deal with them. Third, we present various additional robustness analyses. Lastly, we explore which firms are more likely to launch multiple experiments, and which firms try novel approaches.

6.1 Diversity of Approaches and the Success of Experimentation

We begin by examining, correlationally, the relationships theorized by Propositions 1 and 2 between experimenter scale, the diversity of approaches, and the average success of experimentation. We estimate the following econometric specification:

$$y_{i,t} = \alpha + \beta_1 x_{i,t} + \theta_{i,t} + \gamma_i \times \tau_t + \epsilon_{i,t} \quad (3)$$

Where $y_{i,t}$ represents the market-year level dependent variable *Target Diversity*, or *Share*

of *Pre-Clinical Success*. and $x_{i,t}$ is the average experimenter scale in a therapeutic class–year. We test this specification with and without $\theta_{i,t}$, which is our set of market structure control variables for the number of firms, targets, and pre-clinical experiments within a therapeutic class–year. In all models we include ATC-1 \times Year fixed effects $\gamma_i \times \tau_t$, and standard errors are clustered at the ATC-1 level to account for correlations in errors among market observations within the same industry group.

[Table 5 about here.]

We present the results from this baseline specification in Table 5. Column (1) reports the baseline relationship between experimenter scale and target diversity, showing a positive and significant association. In column (2), we incorporate market structure controls. Notably, the estimated coefficient for β_1 remains consistent across models. Based on the estimate in column (2), our findings suggest that a 1-unit increase in the average experimenter scale (e.g., from an average of 1.1 to 2.1 experiments per firm) corresponds to a 1.65 standard deviation increase in target diversity, which is consistent with Proposition 1.

In columns (3) and (4) we test Proposition 2: small-scale experimenters are more likely to be successful in an individual experiment than large-scale experimenters. Here we regress the *Share of Pre-Clinical Success* on *Average Experimenter Scale*, and find a negative relationship, consistent with our prediction.

Lastly, to test Proposition 3 we estimate the specification detailed in Equation 3 where the independent variable $x_{i,t}$ is *Target Diversity* and the dependent variable $y_{i,t}$ is *At least 1 Pre-Clinical Success*. The results are presented in columns (5) and (6) of Table 5. Consistent with Proposition 3, these results indicate that markets with greater approach diversity in experimentation have a higher probability of at least one success.

[Figure 3 about here.]

6.2 Alternative Explanations

Establishing causality in this context is challenging. The ideal experiment would involve identifying a source of exogenous variation in the cost or value of experimentation for certain

firms in a market. For instance, a reduction in the cost of experimentation for some firms within a market could lead to an increase in the number of experiments they initiate, thereby altering the average experimenter scale without directly affecting the diversity of approaches used. While shocks that influence entire markets do exist—for example, the introduction of Medicare Part D in 2004, which arguably increased the value of experimentation in therapeutic classes with predominantly elderly patients (Dranove et al., 2022)—such shocks tend to affect *all* firms operating in the market.¹⁵

Instead, we take a multipronged approach to evaluate alternative explanations that would undermine our results. Our strongest argument is that neither alternative explanations nor likely sources of unobserved heterogeneity can explain the pattern of empirical results regarding the success of experimentation (Propositions 2 and 3): that greater experimenter scale increases the probability of at least one success but lowers average success. If the relationship between the average experimenter scale and target diversity were driven solely by unobserved common variables, we would not expect higher target diversity to correlate with both a higher probability of at least one success and a lower average share of successes in the market.

For instance, suppose that some markets have more technical opportunities, and this is positively (negatively) associated with the number of approaches observed. Then, there should be a positive (negative) relationship between diversity and average success but also at least one success. There is no reason to expect a relationship between diversity and average size. Alternatively, suppose, for some reason, experimenter-scale is positively related to technological opportunity, which, in turn, is positively related to the diversity of approaches. This would result in a positive association between experimenter-scale and diversity. However, this should also increase, not decrease, average success. Thus, while we cannot establish a causal link between experimenter scale and target diversity, the observed associations between ex-

¹⁵One possible refinement could consider the geographic dimension of drug development. Since Medicare Part D was enacted in the U.S., its impact may have been limited to U.S. pharmaceutical firms within relevant therapeutic classes. However, because pharmaceutical companies operate globally, they are unlikely to be geographically isolated. Moreover, a shock like Medicare Part D could produce mixed effects. While it might encourage firms previously conducting only one experiment to initiate two, it could also incentivize new market entry, enabling firms that otherwise would not have participated to conduct experiments. Consequently, the impact on the composition of the experimenter scale within a market remains ambiguous.

perimeter scale, target diversity, and success outcomes are consistent with the hypothesized causal relationship.

[Table 6 about here.]

There are some possible confounding variables which we can observe and directly control for. The first concerns the discovery of new targets. Our model and simulations assume a fixed pool of approaches from which firms select for experimentation, but the landscape of available approaches may evolve over time. The discovery of new targets and their relationships with therapeutic indications could expand the pool of available approaches, naturally increasing target diversity. Additionally, this dynamic could influence the composition of experimenter scale in the market if single- or multi-experiment firms differ in their likelihood of introducing novel targets.

Second, there could be underlying average firm characteristics within a market that are correlated with both the number of experiments firms' choose to launch and their choice of approach. For instance, younger firms may be less likely to conduct multiple experiments due to limited resources to fund and manage them. Additionally, if younger firms are more inclined to pursue novel approaches, the average firm age in a market could be correlated with target diversity.

In Table 7, we replicate the analysis from Table 5, adding controls for discovery and average firm age. In column (1), the coefficient on *Average Experimenter Scale* increases in magnitude when $\ln(\textit{Discovery})$ is included compared to the baseline result in column (1) of Table 5, suggesting that the baseline model may underestimate the effect due to a downward bias. However, these results are limited by a smaller sample size, stemming from challenges in mapping new target discoveries to therapeutic classes. In column (2), which includes only the control for average firm age, the coefficient remains consistent with the baseline results reported in column (3) of Table 5. Using the estimate from the full model in column (3), we find that a 1-unit increase in the average experimenter scale is associated with a 1.38 standard deviation increase in target diversity, slightly smaller than the baseline result.

As for the average success of experimentation, the inclusion of our control for discovery

suppresses the negative relationship reported between *Average Experimenter Scale* and *Target Diversity*. Still, after controlling for discovery we find no relationship between average experimenter scale and average success. Note that in column (5), which only includes the control for average firm age, the estimated coefficient is indistinguishable from the baseline estimate in column (4) of Table 5.

Lastly, column (7) through (9) report the relationship between diversity and at least one pre-clinical success after controlling for target discovery and average firm age. When $\ln(\textit{Discovery})$ is included, the estimated coefficient on *Target Diversity* is marginally less than the baseline result in column (6) of Table 5. This suggests that the baseline model may slightly overestimate the effect due to an upward bias. Yet, we still find a positive and statistically significant relationship between the diversity of targets and the likelihood of at least one pre-clinical success.

[Table 7 about here.]

[Table 8 about here.]

6.3 Robustness

We also subject our baseline results to a series of robustness tests. The first addresses the complication that arises when a target is known to be a valid approach. In our model, this corresponds to an approach where π —the probability that the approach is a viable solution to the problem—equals one, meaning there is *no* uncertainty about its viability. When $\pi = 1$, the key implication is that outcomes for experiments using the same approach are no longer correlated, as success depends solely on implementation (see Appendix B). In this case, firms will naturally select the approach with which they have the highest probability of implementation success. We restrict the underlying project-level dataset to projects where the target chosen has not been successfully drugged, which we measure by the first launched product. In Appendix Table E.1, we replicate our core results in Table 5, and find quantitatively similar estimates.

Second, we test the sensitivity of our results to alternate time windows. The unit of analysis in our baseline results is the therapeutic class–year. The choice to look within a single year is arbitrary, so we construct two additional datasets, where we define observations within a two-year and five-year window. Table E.2 in the Appendix reports these results. Although we lose the number of observations through a higher level of aggregation, we find results consistent with our preferred specification which defines observations within a 1-year period. The only exception is that we lose statistical significance between average experimenter scale and average success in the therapeutic class–five year period data, and find no relationship instead.

Third, we explore the robustness of our results to alternate measures of key variables. In particular, we test the sensitivity of our results to measuring target diversity with the Herfindahl-Hirschman Index. We also test an alternative measure to success, where we define success as whether a pre-clinical experiment ultimately results in a drug product launch. Table E.3 in the Appendix shows results qualitatively similar to our baseline finding with these alternative measures.

Lastly, we attempt to address possible simultaneity bias by leveraging the panel structure of our data. Specifically, we use the average experimenter scale in the previous year ($i, t - 1$) as an instrument for the average experimenter scale in year t . This instrument relies on the assumption that idiosyncratic factors, unrelated to target diversity or underlying success probabilities, drive some markets to have more firms with lower experimentation costs. If these factors are persistent, the lagged experimenter scale would be correlated with the current experimenter scale but would not capture time-varying unobserved variables that influence experimenter-scale and target diversity.¹⁶ We find, as described in Appendix Table E.4, that the relationship between target diversity and experimenter-scale is strengthened.

¹⁶Using the lagged experimenter scale as an instrument helps address simultaneity bias, but it cannot fully account for unobserved, market-specific and time-persistent factors—such as market dynamics, innovation strategies of firms, or public research funding across markets—that may influence both experimenter scale and target diversity. However, our results are robust to controlling for aggregate industry-level time trends, captured by $ATC-1 \times Year$ fixed-effects (Branstetter et al., 2022).

6.4 Which firms do multiple experiments and introduce novelty?

Our data and results raise at least two questions, which we explore in this section. First, we have not addressed which firms choose to launch multiple experiments. One potential mechanism, demonstrated in our simulations, involves exogenously shifting the share of multi-experiment firms by altering the proportion of firms facing high or low experimentation costs. Alternatively, the same outcome could be achieved by allowing some firms to extract greater value from experiments, prompting them to enter with additional experiments.

A large body of research in Strategy and Economics suggests that firm size likely explains this pattern. Specifically, multi-experiment firms are most likely *large*, as they benefit from lower experimentation costs (Cohen and Klepper, 1996b) and can appropriate greater value from experiments (Arora et al., 2023). This advantage stems, in part, from their complementary downstream co-specialized assets, which are essential for commercialization (e.g., Rosenbloom, 2000; Filippetti and D’Ippolito, 2017).

We use firm ownership as a proxy for firm size, distinguishing between private and public firms. Interestingly, the data provide no evidence that public firms are more likely to conduct multiple experiments in a market. In panel (a) of Figure 4, we show that private firms conduct only one experiment in a therapeutic class–year 61.5% of the time. For public firms, this share is slightly *higher*, at 63.3%. Panel (b) presents the raw counts of experiments by experimenter scale. While public firms conduct significantly more experiments overall, the data reveal that, most of the time, they launch only one pre-clinical experiment in a therapeutic class–year.

Second, considering the discovery of new targets adds an intriguing nuance: which firms are more likely to introduce novelty? Specifically, when a new target–therapeutic class link is discovered, which firms—if any—are more likely to be the first to pursue this approach in pre-clinical trials? In panel (c) of Figure 4, we plot the share of experiments by single-experiment firms and multi-experiment firms (those conducting more than one experiment in a therapeutic class–year) that are the first to use a target in a therapeutic class.

To ensure reliability, we include only projects initiated since 2005, as earlier projects are

more likely to represent the first target–therapeutic class combinations due to the dataset’s starting point in 1995, which excludes targets drugged before this year. Panel (c) shows that single-experiment firms are more likely to drug a novel target: 54.9% of experiments by these firms introduce a new target, compared to 45.9% for multi-experiment firms. However, panel (d) reveals no significant difference between private and public firms in their likelihood of targeting a novel approach.

These results highlight a critical trade-off in the optimal allocation of experiments within a market. Markets dominated by small-scale experimenters are less favorable for innovation in some respects, as they are more likely to rely on a less diverse set of approaches and are less likely to achieve at least one successful experiment. However, these markets are also more likely to introduce new targets, drugging them for the first time.

[Figure 4 about here.]

7 Discussion

This paper addresses the question of how market structure influences the diversity of experimental approaches firms take and, ultimately, the likelihood of solving innovation problems with uncertain solutions. To investigate this, we develop a theoretical model that builds on work by Dasgupta and Maskin (1987) and Bryan and Lemus (2017) that explores how firms choose between approaches under varying market structures, focusing on multi-experiment firms versus single-experiment firms. The model yields three key propositions at the market level. First, firms conducting multiple experiments are more likely to diversify their approaches, whereas many single-experiment firms tend to converge on the most promising approach, reducing market-level diversity. Second, markets dominated by single-experiment firms achieve higher success rates per experiment, due to their focus on the most promising approach. Third, multi-experiment firms maximize the probability of at least one success through diversification, implying that markets with greater experimenter scale are more likely to see at least one experimental success. To test these propositions, we use detailed data on early-stage pharmaceutical drug development, spanning 49,866 projects across 27 years, and

demonstrate how experimenter scale drives approach diversity and influences market-level success.

The findings reveal a robust and significant relationship between experimenter scale and approach diversity. Markets with a higher average experimenter scale—where firms conduct multiple experiments—exhibit a nearly three standard deviation increase in target diversity for every one-unit increase in average scale. This effect is consistent across alternative econometric specifications. Importantly, markets with higher experimenter scale are more likely to achieve at least one successful outcome, with a one-unit increase in scale linked to a 25.9 percentage point rise in the likelihood of at least one experiment progressing to Phase 1 trials. However, this benefit comes at the cost of a lower average success rate per experiment, as larger-scale experimentation dilutes focus on the most promising approaches. Consequently, we also see that markets with greater approach diversity—driven by multi-experiment firms—are more likely to achieve at least one successful outcome, compared to markets dominated by single-experiment firms. These results are robust to alternative explanations, including controls for new target discovery and firm age, underscoring the role of multi-experiment firms in fostering a diverse innovation ecosystem while balancing trade-offs in efficiency and exploratory breadth.

Our findings have implications for the broader literature on experimentation and innovation. By highlighting the dual roles of small- and large-scale experimenters, we add to the understanding of how diversity in approaches arises and evolves within markets (Cohen and Klepper, 1992). Small-scale experimenters act as pioneers, introducing novel approaches that expand the frontier of possibilities, while large-scale experimenters provide depth by thoroughly testing and validating established approaches (Kotha et al., 2011; Ahuja and Morris Lampert, 2001). This complementarity bridges gaps in the literature on technological trajectories, suggesting that innovation ecosystems benefit from balancing exploratory breadth with systematic refinement (Luger et al., 2018; Fang et al., 2010; Chen and Katila, 2008). Our distinction between approach and implementation further informs studies of experimentation by underscoring how shared risks in approach viability can amplify herding behaviors and potentially limit diversity in high-stakes contexts.

For policy, our results suggest a nuanced perspective on fostering innovation through experimentation. Encouraging both small- and large-scale experimenters within markets is crucial to achieving a balance between exploration and exploitation. Policies that support small-scale experimenters, such as grants for early-stage research or incubator programs, can help uncover novel targets and approaches (Bradley et al., 2021). Meanwhile, mechanisms that enable larger firms to scale their experimentation, such as tax incentives or public-private partnerships, ensure that promising but understudied approaches receive the rigorous testing needed for broader application (Dimos and Pugh, 2016; Lerner, 2009). By designing interventions that maintain this balance, policymakers can enhance the robustness of innovation ecosystems and increase the likelihood of market-level breakthroughs.

These insights also have practical implications for the structure of research funding and the design of innovation ecosystems. Allocating resources to create spaces where small-scale experimenters can explore untested ideas while facilitating partnerships with larger organizations can improve innovation outcomes (Cappelen et al., 2012). For instance, collaborative frameworks that combine the agility of startups with the resources of established firms may foster a more diverse and productive experimental landscape (Polidoro Jr and Yang, 2021; Bhaskaran and Krishnan, 2009). Similarly, strategies that mitigate herding, such as incentivizing exploration of high-risk, high-reward approaches, can prevent over-concentration on seemingly safe options, thus ensuring a more resilient innovation pipeline (Von Essen et al., 2020). Together, these findings offer actionable pathways to strengthen the interplay between individual and collective experimentation in driving technological and societal progress.

While our findings provide valuable insights, several limitations warrant discussion. First, the causal interpretation of our results remains challenging. Although we use extensive controls, unobserved factors—such as market-specific shocks to funding, technology, or regulation—may bias our estimates. For example, simultaneous drivers of experimenter scale and diversity, like regulatory changes or scientific breakthroughs, complicate causal attribution. However, such factors are unlikely to account for the full set of empirical associations we document.

Second, our focus on the pharmaceutical industry, while offering rich data, limits general-

izability. Pharmaceuticals involve high R&D costs, long timelines, and significant uncertainty, which may amplify the dynamics we observe. Other industries, with different innovation cycles or competitive pressures, might display distinct patterns. Exploring these relationships in contexts like materials or renewable energy could test the broader applicability of our model and findings. However, we believe this limitation applies primarily to the empirical setting, not the underlying theory. With regard to our theory, our assumptions enforce two key scope conditions: (1) the total value from experimentation is fixed and independent from the *number* of successes, and (2) that there is variation across firms in their ability to conduct multiple experiments. At first, these scope conditions may seem restrictive, but we believe that there are many markets and technologies—such as batteries, quantum computing, and medical devices—which look this way. In these industries, among others, experimentation is costly, which is one reason for variation in experimenter scale. Furthermore, all of these examples share the promise of temporary monopoly rents—that the first success will capture all value for a period of time.

Finally, deeper exploration of mechanisms and firm heterogeneity is needed. We highlight the roles of small- and large-scale experimenters but do not fully unpack how firm size, resources, or market structures influence their contributions to diversity and innovation. Future research should examine how these dynamics vary across industries and firm types, offering richer insights into the optimal design of innovation ecosystems.

By studying how market structure shapes collective experimentation outcomes, we hope this work encourages further analysis of how innovation systems can be designed to solve the complex and pressing problems.

References

- Ahuja, G. and Morris Lampert, C. (2001). Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. *Strategic management journal*, 22(6-7):521–543.
- Arora, A., Cohen, W., Lee, H., and Sebastian, D. (2023). Invention value, inventive capability and the large firm advantage. *Research Policy*, 52(1):104650.
- Ayala-Fontanez, N., Soler, D., and McCormick, T. (2016). Current knowledge on psoriasis and autoimmune diseases. *Psoriasis: Targets and Therapy*.
- Bhaskaran, S. R. and Krishnan, V. (2009). Effort, revenue, and cost sharing mechanisms for collaborative new product development. *Management Science*, 55(7):1152–1169.
- Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 100(5):992–1026.
- Bradley, S. W., Kim, P. H., Klein, P. G., McMullen, J. S., and Wennberg, K. (2021). Policy for innovative entrepreneurship: Institutions, interventions, and societal challenges. *Strategic Entrepreneurship Journal*, 15(2):167–184.
- Branstetter, L., Chatterjee, C., and Higgins, M. J. (2022). Generic competition and the incentives for early-stage pharmaceutical innovation. *Research Policy*, 51(10):104595.
- Bryan, K. A. and Lemus, J. (2017). The direction of innovation. *Journal of Economic Theory*, 172:247–272.
- Camuffo, A., Gambardella, A., Messinese, D., Novelli, E., Paolucci, E., and Spina, C. (2024). A scientific approach to entrepreneurial decision-making: Large-scale replication and extension. *Strategic Management Journal*.
- Cappelen, Å., Raknerud, A., and Rybalka, M. (2012). The effects of r&d tax credits on patenting and innovations. *Research Policy*, 41(2):334–345.
- Chen, B. and Butte, A. (2016). Leveraging big data to transform target selection and drug discovery. *Clinical Pharmacology & Therapeutics*, 99(3):285–297.
- Chen, E. L. and Katila, R. (2008). Rival interpretations of balancing exploration and exploitation: Simultaneous or sequential. *Handbook of technology and innovation management*, 1(1):197.
- Ciarli, T., Kenney, M., Massini, S., and Piscitello, L. (2021). Digital technologies, innovation, and skills: Emerging trajectories and challenges. *Research Policy*, 50(7):104289.
- Cohen, W. M. and Klepper, S. (1992). The tradeoff between firm size and diversity in the pursuit of technological progress. *Small Business Economics*, 4:1–14.
- Cohen, W. M. and Klepper, S. (1996a). Firm size and the nature of innovation within industries: the case of process and product r&d. *The review of Economics and Statistics*, pages 232–243.
- Cohen, W. M. and Klepper, S. (1996b). A reprise of size and r & d. *The Economic Journal*, 106(437):925–951.
- Dara, S., Dhamecherla, S., Jadav, S. S., Babu, C. M., and Ahsan, M. J. (2022). Machine learning in drug discovery: a review. *Artificial Intelligence Review*, 55(3):1947–1999.
- Dasgupta, P. and Maskin, E. (1987). The simple economics of research portfolios. *The Economic Journal*, 97(387):581–595.

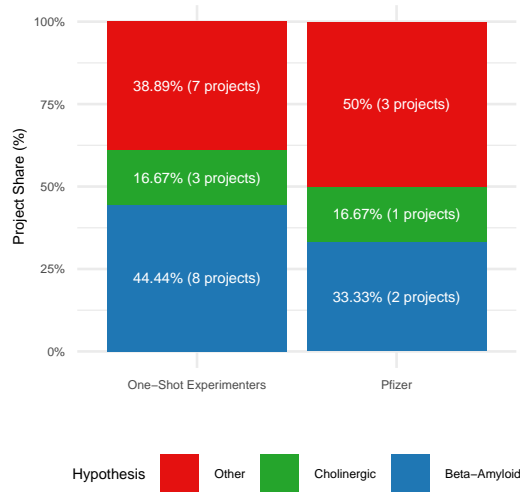
- DiMasi, J. A., Grabowski, H. G., and Hansen, R. W. (2016). Innovation in the pharmaceutical industry: new estimates of r&d costs. *Journal of health economics*, 47:20–33.
- Dimos, C. and Pugh, G. (2016). The effectiveness of r&d subsidies: A meta-regression analysis of the evaluation literature. *Research policy*, 45(4):797–815.
- Dranove, D., Garthwaite, C., and Hermosilla, M. (2022). Does consumer demand pull scientifically novel drug innovation? *The RAND Journal of Economics*, 53(3):590–638.
- Fang, C., Lee, J., and Schilling, M. A. (2010). Balancing exploration and exploitation through structural design: The isolation of subgroups and organizational learning. *Organization Science*, 21(3):625–642.
- Filippetti, A. and D’Ippolito, B. (2017). Appropriability of design innovation across organisational boundaries: exploring collaborative relationships between manufacturing firms and designers in italy. *Industry and Innovation*, 24(6):613–632.
- Frankel, A. P., Krieger, J. L., Li, D., and Papanikolaou, D. (2023). Evaluation and learning in r&d investment.
- Gans, J. S., Stern, S., and Wu, J. (2019). Foundations of entrepreneurial strategy. *Strategic Management Journal*, 40(5):736–756.
- Gross, D. P. and Sampat, B. N. (2023). The world war ii crisis innovation model: What was it, and where does it apply? *Research Policy*, 52(9):104845.
- Guo, J., Zhang, H., Lin, W., Lu, L., Su, J., and Chen, X. (2023). Signaling pathways and targeted therapies for psoriasis. *Signal Transduction and Targeted Therapy*, 8(1):437.
- Hong, L. and Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*, 101(46):16385–16389.
- Hughes, J. P., Rees, S., Kalindjian, S. B., and Philpott, K. L. (2011). Principles of early drug discovery. *British journal of pharmacology*, 162(6):1239–1249.
- Kang, H., Fu, L., Funk, R. J., Wang, X., Ding, J., Liang, S., Wang, J., Zhou, L., and Zhou, C. (2024). Scientific and technological knowledge grows linearly over time. *arXiv preprint arXiv:2409.08349*.
- Knowles, J. and Gromo, G. (2003). Target selection in drug discovery. *Nature Reviews Drug Discovery*, 2(1):63–69.
- Koning, R., Hasan, S., and Chatterji, A. (2022). Experimentation and start-up performance: Evidence from a/b testing. *Management Science*, 68(9):6434–6453.
- Kotha, R., Zheng, Y., and George, G. (2011). Entry into new niches: The effects of firm age and the expansion of technological capabilities on innovative output and impact. *Strategic management journal*, 32(9):1011–1024.
- Krieger, J., Li, D., and Papanikolaou, D. (2022). Missing novelty in drug development. *The Review of Financial Studies*, 35(2):636–679.
- Krieger, J. L. (2021). Trials and terminations: Learning from competitors’ r&d failures. *Management Science*, 67(9):5525–5548.
- Lerner, J. (2009). Boulevard of broken dreams: why public efforts to boost entrepreneurship and venture capital have failed—and what to do about it.

- Lieberman, M. B. and Asaba, S. (2006). Why do firms imitate each other? *Academy of management review*, 31(2):366–385.
- Liu, P.-P., Xie, Y., Meng, X.-Y., and Kang, J.-S. (2019). History and progress of hypotheses and clinical trials for alzheimer’s disease. *Signal transduction and targeted therapy*, 4(1):29.
- Luger, J., Raisch, S., and Schimmer, M. (2018). Dynamic balancing of exploration and exploitation: The contingent benefits of ambidexterity. *Organization science*, 29(3):449–470.
- Margalef, R. (1958). Information theory in ecology. *General Systems*, 3:36–71.
- Nelson, A., Anthony, C., and Tripsas, M. (2023). “if i could turn back time”: Occupational dynamics, technology trajectories, and the reemergence of the analog music synthesizer. *Administrative Science Quarterly*, 68(2):551–599.
- Nelson, R. R. (1961). Uncertainty, learning, and the economics of parallel research and development efforts. *The review of economics and statistics*, pages 351–364.
- Nestle, F. O., Kaplan, D. H., and Barker, J. (2009). Psoriasis. *New England Journal of Medicine*, 361(5):496–509.
- Piller, C. (2022). Blots on a field? *Science (New York, NY)*, 377(6604):358–363.
- Piller, C. (2024). Researchers plan to retract landmark alzheimer’s paper containing doctored images. *Science*.
- Polidoro Jr, F. and Yang, W. (2021). Corporate investment relationships and the search for innovations: An examination of startups’ search shift toward incumbents. *Organization Science*, 32(4):909–939.
- Prager, K. (2021). Seek diversity to solve complexity. *Nature*.
- Rosenbloom, R. S. (2000). Leadership, capabilities, and technological change: The transformation of ncr in the electronic era. *Strategic management journal*, 21(10-11):1083–1103.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423.
- Sorenson, O. (2024). Theory, search, and learning. *Strategy Science*.
- Tan, D. (2023). The road not taken: Technological uncertainty and the evaluation of innovations. *Organization Science*, 34(1):156–175.
- Thomke, S., Von Hippel, E., and Franke, R. (1998). Modes of experimentation: an innovation process—and competitive—variable. *Research Policy*, 27(3):315–332.
- Tranchoero, M. (2023). Finding diamonds in the rough: Data-driven opportunities and pharmaceutical innovation. *Working Paper. University of California, Berkeley*.
- Von Essen, E., Huysentruyt, M., and Miettinen, T. (2020). Exploration in teams and the encouragement effect: Theory and experimental evidence. *Management science*, 66(12):5861–5885.

FIGURE 1: A 2x2 FRAMEWORK OF EXPERIMENTATION: APPROACHES VS. IMPLEMENTATION

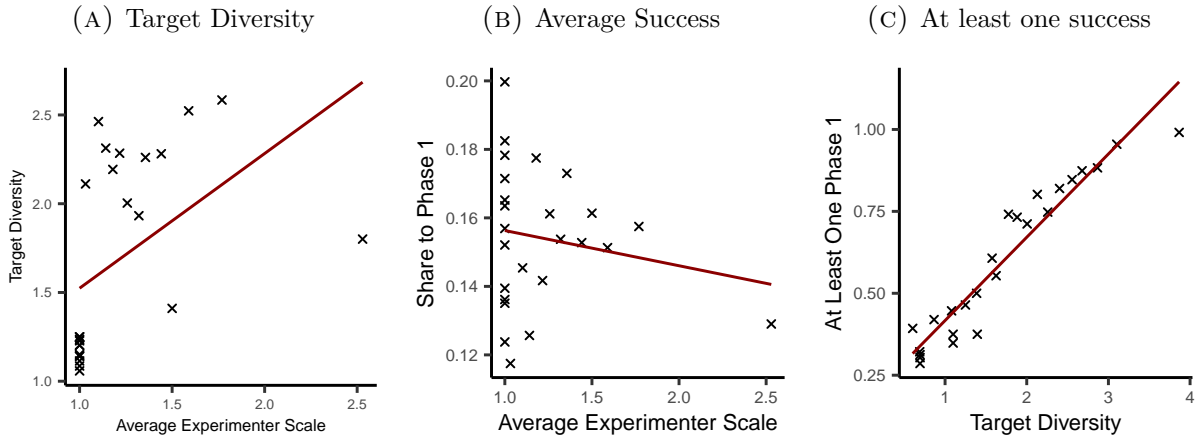
		IMPLEMENTATION	
		Success	Failure
APPROACH	Correct	<p>The approach was correct, and the implementation confirmed it.</p> <p>Examples: using CRISPR to edit specific genes; the hypothesis of the Higgs field was confirmed by experiments that implemented the approach at the Large Hadron Collider.</p>	<p>The underlying approach was correct, but the specific implementation failed.</p> <p>Example: the Wright brother's initial attempts at heavier-than-air flight</p>
	Wrong	<p>The approach was wrong, but the implementation attempt led to an unexpected, serendipitous discovery.</p> <p>Examples: The discovery of penicillin; viagra for erectile (originally developed for hypertension).</p>	<p>The approach was incorrect, and the implementation confirmed this by failing.</p> <p>Example: producing a better vacuum to fix the problem of blackening tungsten filament light bulbs.</p>

FIGURE 2: AD PROJECTS STARTED BY HYPOTHESIS FOR PFIZER AND SINGLE-SHOT FIRMS 2007-2008



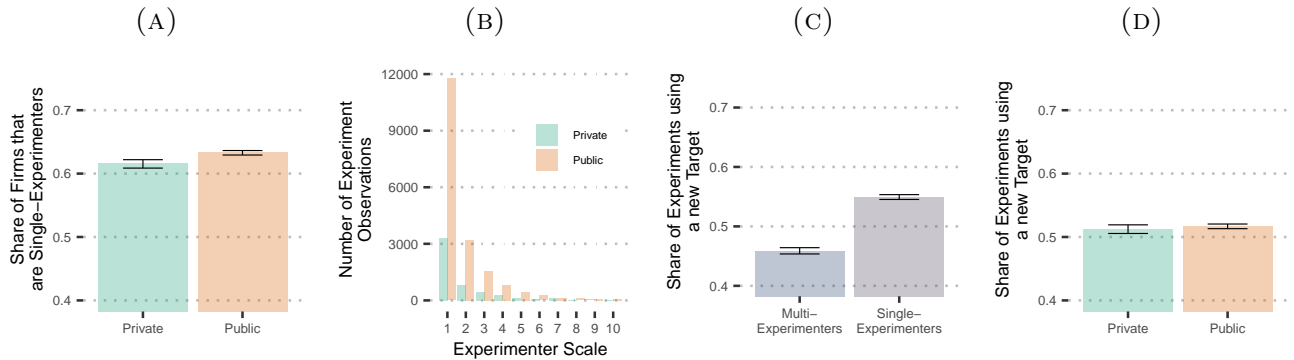
Notes: This Figure compares the allocation of hypotheses across projects started by Pfizer to projects started by single-experimenter firms in 2007 and 2008. Single-experimenter firms are those which started only one AD drug development project in the period 2007-2008. We include projects where we can identify the year in which preclinical development started and the target used.

FIGURE 3: BINNED SCATTER PLOTS OF THE MARKET-LEVEL RELATIONSHIP BETWEEN EXPERIMENTER SCALE, THE DIVERSITY OF APPROACHES, AND SUCCESS



Notes: This Figure shows non-parametric binned scatter plots of the relationship between *Target Diversity*, *Average Experimenter Scale*, the *Share of Pre-Clinical Success*, and *At least 1 Pre-Clinical*, corresponding to the parametric estimates in Table 5. Each plot uses 25 bins.

FIGURE 4: THE RELATIONSHIPS BETWEEN OWNERSHIP, EXPERIMENTER SCALE, AND THE LIKELIHOOD OF EXPERIMENTING WITH NOVEL TARGETS



Notes: This figure explores the non-parametric relationships between ownership, experimenter scale, and the likelihood of experimenting with novel targets. The data for all panels only includes projects from 2005 and onwards. This is to address the concern that novelty may be artificially inflated in earlier years because our data starts in 1995, and novelty is defined with respect to what came before. Panel (a) is a bar chart comparing the average share of private and public firms that conduct multiple experiments in a therapeutic class-year. Panel (b) plots the count of experiments conducted by experimenter scale for both public and private firms. Panel (c) compares the share of experiments conducted by multi-experiment and single-experiment firms that use a novel target. Panel (d) compares the share of experiments conducted by private and public firms that use a novel target.

TABLE 1: DRUG DEVELOPMENT PROJECTS STARTED FOR ALZHEIMER’S DISEASE BY HYPOTHESIS 1998-2008

Hypothesis	Number of Projects	Project Share (%)
Beta-Amyloid	49	39.2
Cholinergic	22	17.6
Other	54	43.2

Notes: This table shows the distribution of drug development projects in Alzheimer’s Disease between 1998 and 2008 by hypothesis. We include projects where we can identify the year in which preclinical development started and the target used. Our data describe the target that a drug is intended to act upon, not the broader underlying theory or hypothesis. For illustration purposes, in this table, we group targets into broader hypothesis groups. The beta-amyloid hypothesis includes drugs that target amyloid-beta, beta-secretase, adrenoceptor beta 1 and 2, and glutamate metabotropic. The cholinergic hypothesis includes drug projects that target cholinergic receptors and acetylcholine. Other hypotheses for AD include the tau hypothesis, mitochondrial cascade hypothesis, calcium homeostasis hypothesis, neurovascular hypothesis, inflammatory hypothesis, metal ion hypothesis, and lymphatic system hypothesis (Liu et al., 2019).

TABLE 2: SUMMARY OF PROPOSITIONS

1	Holding the number of experiments constant, a firm pursuing multiple experiments will explore a greater variety of approaches compared to an equivalent number of projects, each pursued by a separate firm.
2	Small-scale experimenters are more likely to be successful in an individual experiment than large-scale experimenters.
3	Markets with greater approach diversity in experimentation have a higher probability of at least one success.

TABLE 3: DESCRIPTIVE STATISTICS

	Mean	St. Dev.	Min	Q1	Median	Q3	Max	N
Target Diversity	1.776	0.857	0.377	1.099	1.609	2.369	5.144	2523
Average Experimenter Scale	1.225	0.374	1.000	1.000	1.091	1.333	6.000	2523
Average Experimenter Scale (t-1)	1.225	0.370	1.000	1.000	1.100	1.333	6.000	2384
Multi-Experiment Share	0.138	0.170	0.000	0.000	0.079	0.250	1.000	2523
Multi-Experiment Share (t-1)	0.138	0.168	0.000	0.000	0.083	0.250	1.000	2384
ln(Discovery)	3.335	1.834	0.000	2.079	3.401	4.663	7.675	809
Average Firm Age	37.286	30.866	0.000	13.500	28.545	52.778	162.000	2447
At least 1 Pre-Clinical Success	0.621	0.485	0.000	0.000	1.000	1.000	1.000	2523
Share of Pre-Clinical Success	0.155	0.182	0.000	0.000	0.111	0.250	1.000	2523
At least 1 Drug Launch	0.170	0.376	0.000	0.000	0.000	0.000	1.000	2523
Share of Drug Launch	0.028	0.086	0.000	0.000	0.000	0.000	1.000	2523
Project Start Year	2010.126	7.879	1996.000	2004.000	2010.000	2017.000	2023.000	2523
Number of Firms	8.897	12.276	2.000	3.000	5.000	10.000	165.000	2523
Number of Projects Started	11.568	19.329	2.000	3.000	6.000	13.000	306.000	2523
Number of Targets	9.782	14.772	2.000	3.000	5.000	11.000	210.000	2523

Notes: This table presents descriptive statistics for all variables in our therapeutic class-year (i, t) dataset, of which there are $N = 2,523$ observations. For some variables, there are fewer observations and/or there are missing data. The lagged measures of *Average Experimenter Scale (t - 1)* and *Multi-Experimenter Share (t - 1)* have 2,384 observations, as we do not have data for the first year in our sample (1996). Only 809 observations have values for $\ln(\text{Discovery})$ due to the matching required between Pharmaprojects and GWAS (See Appendix D.4 for more details). Similarly, *Average Firm Age* was calculated for the 2,447 observations where at least one firm was matched to Pitchbook.

TABLE 4: CORRELATION MATRIX

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Target Diversity	1.00														
Average Experimenter Scale	0.27	1.00													
Average Experimenter Scale (t-1)	0.22	0.26	1.00												
Multi-Experiment Share	0.26	0.82	0.22	1.00											
Multi-Experiment Share (t-1)	0.17	0.21	0.81	0.20	1.00										
ln(Discovery)	0.24	0.02	-0.02	-0.01	-0.02	1.00									
Average Firm Age	-0.05	0.03	0.02	0.05	0.03	-0.15	1.00								
At least 1 Pre-Clinical Success	0.46	0.19	0.13	0.16	0.11	0.12	-0.04	1.00							
Share of Pre-Clinical Success	-0.03	0.06	0.06	0.06	0.08	-0.04	-0.08	0.62	1.00						
At least 1 Drug Launch	0.22	0.11	0.15	0.12	0.08	-0.05	0.07	0.15	0.09	1.00					
Share of Drug Launch	-0.10	-0.00	0.01	0.04	-0.03	-0.08	0.07	0.03	0.12	0.67	1.00				
Project Start Year	-0.02	-0.06	-0.04	-0.09	-0.05	0.46	-0.35	-0.07	-0.04	-0.26	-0.15	1.00			
Number of Firms	0.79	0.30	0.29	0.26	0.24	0.22	-0.10	0.30	0.02	0.18	-0.06	0.08	1.00		
Number of Projects Started	0.73	0.37	0.32	0.31	0.26	0.19	-0.09	0.27	0.04	0.17	-0.05	0.07	0.98	1.00	
Number of Targets	0.77	0.34	0.31	0.29	0.25	0.20	-0.09	0.28	0.03	0.18	-0.06	0.07	0.99	0.99	1.00

Notes: This table details the correlation coefficient between all variables in our therapeutic class-year (i, t) dataset, of which there are $N = 2,523$ observations. For some variables, there are fewer observations and/or there are missing data. The lagged measures of *Average Experimenter Scale (t - 1)* and *Multi-Experimenter Share (t - 1)* have 2,384 observations, as we do not have data for the first year in our sample (1996). Only 809 observations have values for $\ln(\text{Discovery})$ due to the matching required between Pharmaprojects and GWAS (See Appendix D.4 for more details). Similarly, *Average Firm Age* was calculated for the 2,447 observations where at least one firm was matched to Pitchbook.

TABLE 5: THE MARKET-LEVEL RELATIONSHIP BETWEEN EXPERIMENTER SCALE, THE DIVERSITY OF APPROACHES, AND SUCCESS

	Target Diversity		Share of Pre-Clinical Success		At least 1 Pre-Clinical Success	
	(1)	(2)	(3)	(4)	(5)	(6)
Average Experimenter Scale	0.682*** (0.188)	0.651*** (0.093)	-0.027*** (0.004)	-0.029*** (0.005)		
Target Diversity					0.239*** (0.024)	0.330*** (0.024)
Market Structure Controls	No	Yes	No	Yes	No	Yes
ATC-1×Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,503	2,503	2,503	2,503	2,503	2,503
Adj R-squared	0.124	0.754	0.063	0.063	0.200	0.219

Notes: This Table reports baseline regression results testing each Proposition. In columns (1) and (2), we test Proposition 1 by regressing *Target Diversity* on *Average Experimenter Scale*, where the difference between the two models is that column (2) includes the set of *Market Structure Controls*. *Target Diversity* is the Shannon entropy of the relative abundance of targets employed in pre-clinical experiments, and *Average Experimenter Scale* is the average number of pre-clinical experiments started by firms in a therapeutic class-year. In columns (3) and (4), we test Proposition 2 by regressing the *Share of Pre-Clinical Success* on *Average Experimenter Scale*. Lastly, columns (5) and (6) correspond to Proposition 3, where we explore the relationship between *Target Diversity* and *At least 1 Pre-Clinical Success*. All models are estimated with OLS. Robust standard errors clustered at the ATC-1 level are shown in parentheses. Significance codes: * p<.1, ** p<.05, *** p<.01

TABLE 6: LINKING APPROACH DIVERSITY TO THE SUCCESS OF EXPERIMENTS

	Share of Pre-Clinical Success		At least 1 Pre-Clinical Success	
	(1)	(2)	(3)	(4)
Target Diversity	-0.005 (0.011)	0.005 (0.014)	0.330*** (0.024)	0.334*** (0.031)
Average Experimenter Scale		-0.032*** (0.010)		-0.015 (0.033)
Market Structure Controls	Yes	Yes	Yes	Yes
ATC-1×Year FE	Yes	Yes	Yes	Yes
Observations	2,503	2,503	2,503	2,503
Adj R-squared	0.060	0.062	0.219	0.219

Notes: This Table reports regression results that link success outcomes to the diversity of targets, while controlling for the average scale of experimenters in the market. In column (1), we regress *Share of Pre-Clinical Success* on *Target Diversity*, and in column (2) we include *Average Experimenter Scale* as a control. Columns (3) and (4) are similar, but the dependent variable is *At least 1 Pre-Clinical Success*. All models are estimated with OLS. Robust standard errors clustered at the ATC-1 level are shown in parentheses. Significance codes: * p<.1, ** p<.05, *** p<.01

TABLE 7: THE RELATIONSHIP BETWEEN EXPERIMENTER SCALE AND TARGET DIVERSITY WHILE CONTROLLING FOR TARGET DISCOVERY AND FIRM AGE

	Target Diversity			Share of Pre-Clinical Success			At least 1 Pre-Clinical Success		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Average Experimenter Scale	0.724*** (0.092)	0.678*** (0.087)	0.718*** (0.089)	0.029 (0.018)	-0.024*** (0.005)	0.029 (0.018)			
Target Diversity							0.309*** (0.020)	0.330*** (0.023)	0.301*** (0.018)
ln(Discovery)	0.032** (0.013)		0.030** (0.013)	0.000 (0.006)		0.000 (0.006)	0.006 (0.015)		0.006 (0.015)
Average Firm Age		-0.000 (0.000)	-0.001 (0.001)		-0.000 (0.000)	-0.001*** (0.000)		-0.001 (0.000)	-0.001* (0.001)
Market Structure Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ATC-1×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	762	2,427	756	762	2,427	756	762	2,427	756
Adj R-squared	0.770	0.755	0.771	-0.007	0.063	0.001	0.210	0.217	0.205

Notes: This Table reports baseline regression results testing each Proposition while controlling for two observable potential confounders. In columns (1) through (3), we regress *Target Diversity* on *Average Experimenter Scale*. Column (1) includes $\ln(\text{Discovery})$, column (2) includes *Average Firm Age*, and column (3) includes both variables. Recall that $\ln(\text{Discovery})$ is the natural logarithm of the count of publications in GWAS that report a novel target, and *Average Firm Age* is the average age of distinct experimenters (firms) in a therapeutic class-year. The remaining models follow the same pattern, where the dependent variable in columns (4) through (6) is *Share of Pre-Clinical Success*, and in columns (7) through (9) the dependent variable is *At least 1 Pre-Clinical Success*. Note that in the final three columns—consistent with the baseline results in Table 5 that test Proposition 3—the independent variable of interest is *Target Diversity*. All models are estimated with OLS. Robust standard errors clustered at the ATC-1 level are shown in parentheses. Significance codes: * $p < .1$, ** $p < .05$, *** $p < .01$

TABLE 8: ALTERNATIVE EXPLANATIONS

Alternative Explanation	Argument Against
If some markets have more technical opportunities, and this is positively (negatively) associated with the number of approaches observed, then there should be a positive (negative) relationship between diversity and average success but also at least one success. There is no reason to expect a relationship between diversity and average size.	Results in Table 6 indicate a positive relationship between target diversity and at least one success, but no relationship between target diversity and the share of success
If experimenter-scale is positively related to technological opportunity, which, in turn, is positively related to the diversity of approaches, this would result in a positive association between experimenter-scale and diversity.	However, this should also increase, not decrease, average success, as Table 5 shows.
The discovery of new targets and their relationships with therapeutic indications could expand the pool of available approaches, naturally increasing target diversity. Additionally, this dynamic could influence the composition of experimenter scale in the market if single- or multi-experiment firms differ in their likelihood of introducing novel targets.	Table 7 shows that our baseline results from Table 5 is robust to the inclusion of our control for target discovery.
Younger firms may be less likely to conduct multiple experiments due to limited resources to fund and manage them. Additionally, if younger firms are more inclined to pursue novel approaches, the average firm age in a market could be correlated with target diversity.	Table 7 shows that our baseline results from Table 5 is robust to the inclusion of our control for average firm age.