Charles A. Dice Center for Research in Financial Economics

From Anecdotes to Insights: Streamlining the Research Idea Generation Process

Itzhak Ben-David,
The Ohio State University and NBER

fisher.osu.edu

From Anecdotes to Insights: Streamlining the Research Idea Generation Process

Itzhak Ben-David[*]

Monday 7th October, 2024

Abstract

This paper explores strategies for generating and evaluating novel research ideas. Researchers can identify promising ideas by systematically exposing themselves to new, practitioner-relevant information and by contrasting emerging facts with existing theories. Additionally, by identifying the necessary conditions that are required for an idea to become a viable research project, researchers can quickly discard low-prospect ideas, freeing up mental space and time to evaluate new research opportunities.

Keywords: Research Methods, Idea Generation, Academic Research, Research Process, Research Strategies, Research Productivity
JEL Classification: A11, A20, B40, G00, G10

1 Generate Fast, Fail Fast: Maximizing Research Throughput

Finding new research ideas can be daunting, especially for researchers early in their careers. Nothing is more frustrating than sifting through the well-written papers in *The Journal of Finance*, hoping to get inspired by one of them and that a publication-worthy idea will emerge. This moment is exceptionally rare.

The challenge of finding the golden research idea is most acute at the start of one's career. Getting your research on track early, especially with a short tenure clock (e.g., six years), is critical. In my experience, only one in ten ideas that initially seem promising actually becomes a paper published in an A-tier journal. Why? Because the process is a funnel with a low convergence rate. One in five ideas will end up as a working paper, and only half will be published in a top-tier journal. Moreover, the time I spend on home-run ideas is far less than the time spent on poor ideas. The reason is that a home-run idea will go through the publication relatively fast while a poor idea will bounce between journals and could take many years to find a home. So, to build your publication record, you must increase the throughput of robust ideas and minimize the time spent on poor ideas that lead to unsuccessful papers.

These statistics raise two immediate questions: First, how does one generate the initial flow of ideas? Second, even with a steady flow of ideas, how can they be efficiently screened, given that processing each could take months: writing a literature review, gathering and cleaning data, analyzing the data, and drafting a short paper?

These questions highlight the need for an efficient approach to generating and evaluating ideas. I'll share the process that works for me, hoping it provides helpful guidance.

Before diving in, a quick disclaimer: Early-career researchers receive a lot of advice on finding ideas, such as focusing on industry-relevant topics (Lowry, 2024), or staying up-to date with ongoing research (Weisbach, 2021). It is certainly the case that different approaches

could work for different people, so early-career researchers should experiment and find out what suits them best. What I describe here is simply what works for me.

2   Strategies for Idea Generation in Academia

2.1   What Makes an Idea Worth Pursuing?

Let's first consider what constitutes an interesting and impactful idea. Generally, academic theories aim to model the world, i.e., reveal its laws of motion. Imagine the state of knowledge as a vector space. We can visualize reality as a point in this multidimensional vector space, though its exact location is unobserved. The current state of academic literature, representing our best theory of how the world works, is another point in this space.

In some fields, scientific theories closely align with reality, meaning their distance in our vector space is minimal. For example, Newton's Laws explain kinetics well, as the theory generates accurate predictions about the movement of physical bodies. Here, the "world" and the "literature" are very close in the vector space, leaving little room for innovation in this area of physics. However, this was not always the case. Figure 1, Panel (a), illustrates different theories of physics within the vector space. If the field of interest is sub-atomic particles, then Newtonian physics is far from reality: it does not provide accurate predictions.

We should all consider ourselves fortunate, as there is broad agreement among economists that the state of the in financial economics is still far from a satisfactory description of reality (Caballero, 2010; Graham, 2022). The world is full of puzzles and mysteries, they say. This lack of understanding (or agreement) of how the world works is our opportunity to write great papers.

A interesting and impactful research idea bridges this gap between the current state of the literature and the true, albeit unobservable, state of the world that the academic community is interested in understanding. In our vector representation, the best research idea one could generate can be visualized as an arrow running from the current literature to the true state

Figure 1: **Mapping Models and Reality in the Vector Space**

Panel (a) illustrates physics models and reality in a vector space. Panel (b) illustrates a model, reality, and the ideal paper in a vector space.

(a) World and theories in physics

(b) What is a "Good Idea"?
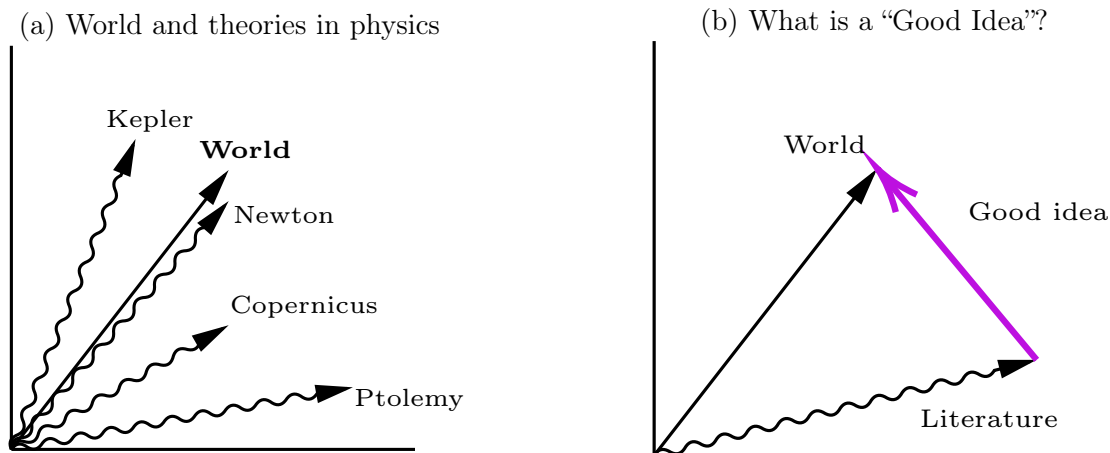
of the world, as shown in Figure 1, Panel (b). The figure illustrates that a research idea is more important and impactful when there is a significant difference between the academic consensus and reality. In other words, good research ideas highlight the dissonance between the academic community's view of the world and the world's true state.

2.2 Sources of New Research Ideas

Since the ideal research idea bridges the gap between the current state of the literature and reality, we can use the vector space representation to guide our search for this idea. Here are some approaches, though not all are equally productive.

Start with the literature. As a PhD student, I was advised to read existing literature to find ideas. It was the post-Enron/Worldcom scandals era, and I wanted to write about the hottest topic du jour: corporate governance (akin to ESG or AI today...). I spent days reading the latest finance journals and looking for an angle for a new paper. Nothing came out of this process. Now I realize that the reason for lack of success was that using the current literature as an anchor can generate only minor marginal extensions in most cases.

So, starting from the literature is a no-go for me. Figure 2 illustrates this point in our vector space. Panel (a) shows the potential extensions of the literature. But how do we know which direction leads closer to reality? Panel (b) shows that the real world could be very far from where the literature currently stands, so minor extensions will likely be insignificant and potentially even pulling away from where reality is.

Figure 2: **Looking for Research Ideas in the Existing Literature**

Panel (a) shows the directions that a researcher could take when looking for research ideas in the current literature. Panel (b) shows that the real world could be very distant from the current literature.

(a) New ideas extending the literature

(b) ...but literature is far from reality
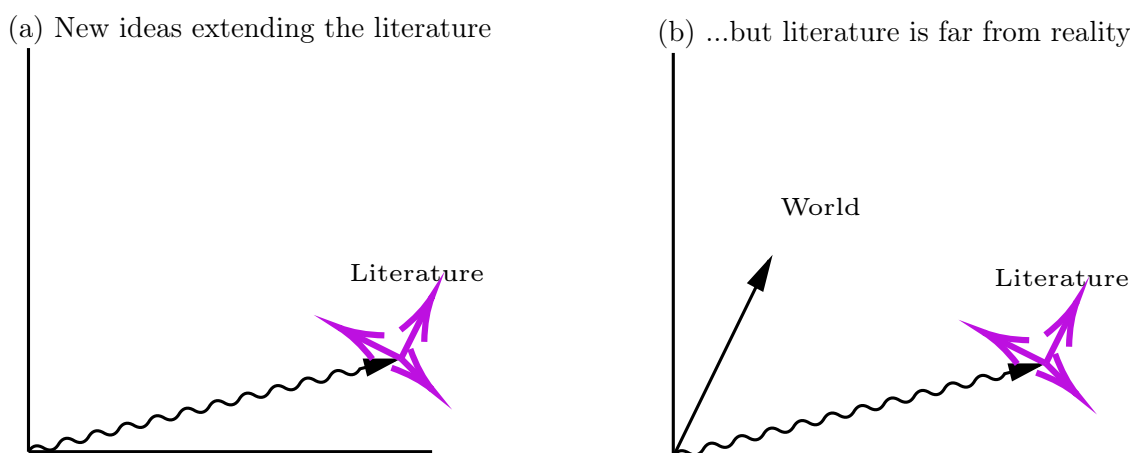
That said, it's essential to have a solid understanding of the foundational theories in your field, such as Modigliani and Miller's propositions in finance. Reading the literature alone is unlikely to spark truly innovative ideas. However, this knowledge is critical for identifying where the literature stands and for challenging existing views. Without such understanding, it is difficult to assess whether reality is near or far from the current literature.

Start with a cool dataset. Many students believe that finding the right dataset is the key to a publishable paper. However, a dataset is not a research idea. While numerous research questions can be answered with any dataset, most of these questions won't interest financial economists and, therefore, won't be publishable. For instance, consider a dataset of millions of credit card statements over a decade. It might seem ideal for studying household finance, but which "reality" are we aiming to uncover? Are we looking at behavioral biases

4

in spending patterns, household sensitivity to retailer promotions, online versus brick-and-mortar shopping habits, or coffee-drinking trends?

The abundance of potential questions stemming from a new dataset can actually hinder the research process. Without a specific research question in mind, it's tempting to spend precious time downloading, cleaning, and coding the data. As a student, I actually loved tinkering with new datasets. I didn't get a job market paper out of it, but loved every moment. In retrospect, it wasn't a productive activity.

Figure 3, Panel (a), illustrates the wealth of ideas that a new dataset can ignite. Panel (b) presents the downside of using a new dataset as a starting point: there is no "hook"—it is not clear which idea is relevant because it is not clear which part of our understanding of the world it can help shed light on. A good research idea needs to bridge a gap between the current literature and a reality that academics are interested in.

Figure 3: **Starting Your Research via a New Dataset**

Panel (a) illustrates potential ideas that one could develop from a new dataset. Panel (b) illustrates the difficulty in linking the research ideas to specific realities that would result in publication-worthy papers.

(a) Ideas coming out of a dataset

(b) ...but which reality do they explore?
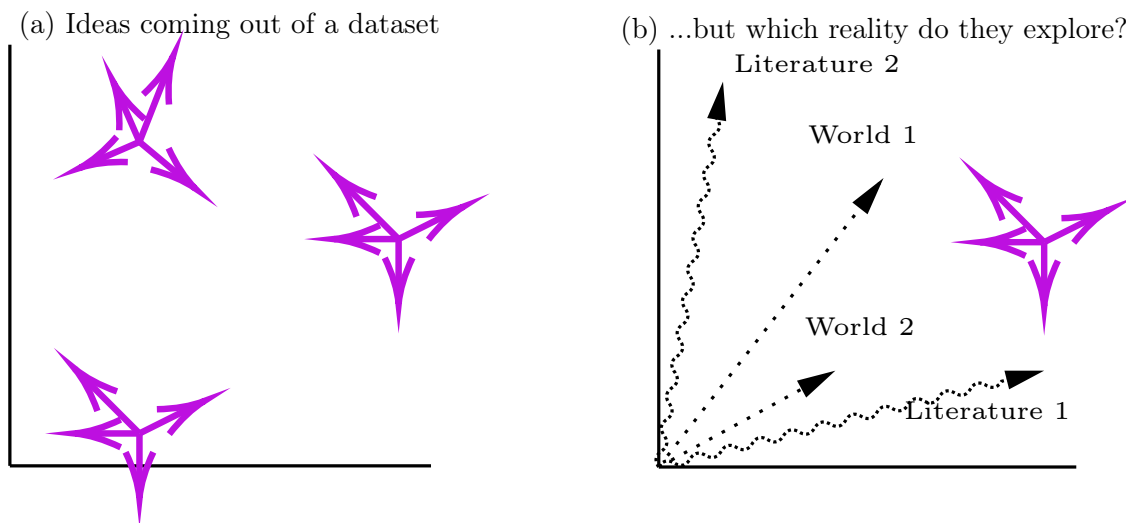
Start with anecdotes. I've long believed that to understand the world, we must be attuned to it. This means actively collecting and mentally cataloging anecdotes. Over time, these anecdotes accumulate and reveal patterns. A researcher should constantly ask, "Is the

way people behave consistent with how economists think people behave?" A good research idea bridges the gap between what economists currently believe and what the world truly is.

This approach is illustrated in Figure 4. Rather than starting with the literature or available data, begin by observing the behavior of real actors: households, managers, and investors. These anecdotes provide clues about where reality lies (Panel (a)). Then, ask yourself whether the literature is close to or far from where reality appears to be based on the anecdotes (Panel (b)). The arrows connecting the current academic consensus to reality represent your ideal paper. There are multiple arrows in Panel (b) because we never know for sure where reality is; the best we can do is bridge the literature to where we *hypothesize* reality is, depending on the anecdotes we manage to collect.

Figure 4: **Finding Research Ideas via Anecdotes**

Panel (a) shows how anecdotes can signal where the relevant reality is in a vector space. Panel (b) shows how one can contrast the academic status quo with the anecdotes to generate a research idea.

(a) Anecdotes mark where reality is

(b) ...and the literature is often far
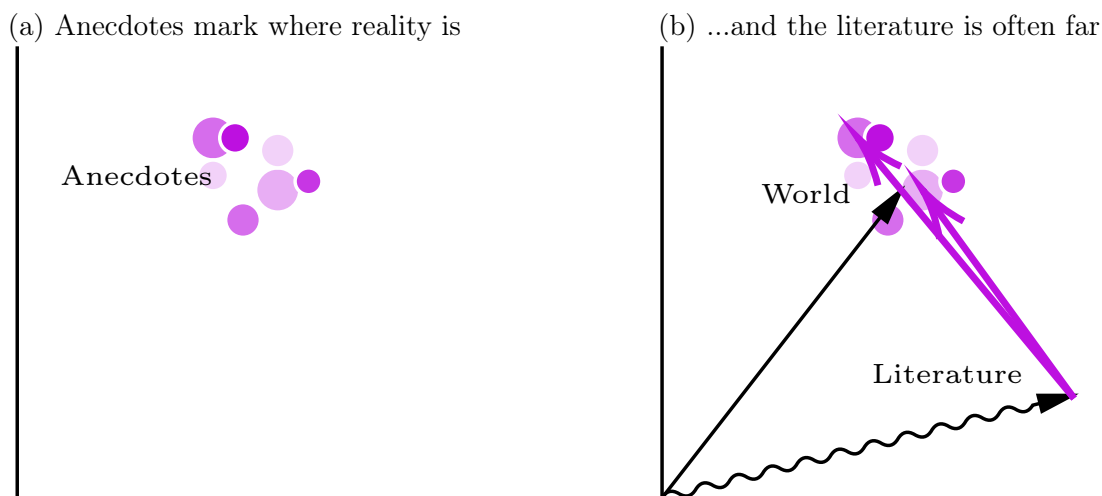
Where there's a significant gap between how economists perceive the world and how real people experience it, there's potential for a new research project. Your next goal is to ask, "How do I present compelling evidence that, while economists think people behave like X, in reality, they behave like Y?" From here, you can start building your case. Many uncertainties remain, and the chance that this will become your job market paper is still low, but now you have a potential research question to work with.

A research idea worthy of your time is one that brings something new to the table, i.e., it proposes that readers should update their priors. For example, your paper could introduce new evidence for an untested theory or provide a better quantification of an important coefficient in another theory. As such, these projects push our literature closer to the true state of reality.

My research ideas appear in my mind in many settings: when I read the news, meet old friends for a beer, or chat with new acquaintances at a dinner party. I'm constantly attentive, trying to make sense of the world around me. In the back of my mind, I'm always asking, "Is what I'm observing consistent with how economists think about the world?" When I notice a discrepancy between the way actors—investors, employees, or managers—describe their decision-making processes and how economists believe these people think and act, I see potential for a research paper. This is the gap between the literature and reality depicted in Figure 1, Panel (b).

2.3 Example: NPV or EPS?

Here's an example to illustrate my point. There is a clear discrepancy between what we teach in corporate finance classes and how managers describe their decision-making processes. I noticed a gap between what I teach my students and what my practitioner guest speakers say in the classroom. In the classroom and academic papers, we assume that the managers' objective is to maximize net present value (NPV). We teach our students that the main reason to borrow is to benefit from the interest tax shield. However, my guest speakers often talk about borrowing to increase the return on their investment. They consider debt as cheap capital, especially when interest rates are low. Tax shield considerations do appear to be on their radar.

A similar disconnect appears in publicly traded firms. When discussing corporate policies, managers and investors seem focused on earnings per share (EPS). NPV and EPS are not the same constructs, revealing a significant dissonance between how financial economists

7

understand corporate decision-making and how managers actually think. Since the 1960s, academic papers have almost exclusively started their exploration of corporate policies by assuming that managers maximize NPV. This approach has led the field to a roadblock: NPV alone cannot explain the corporate policies we observe, leading researchers to propose complex explanations for each empirical 'puzzle.'

However, managers don't evaluate their decisions in NPV terms, nor do they view their corporate policies as tools to overcome the complications that economists suggest they do. In earnings calls, the main discussion topic between managers, analysts, and investors is last period's EPS and its expected growth over the coming period. When announcing acquisitions or share buybacks, managers often explain that these actions will boost their firm's EPS. Could it be that financial economists misinterpret the role of corporate policies and that these policies, at least in part, are targeted to maximize EPS? If EPS maximization rather than NPV determines corporate policies, this mechanism could help bridge the gap between the current state of the literature and reality. It certainly seems so (Ben-David and Chinco, 2023).

The NPV vs EPS example is also telling about how ideas can and cannot be sourced. If one thinks about finding impactful research ideas as extensions of the current academic literature, there is little chance of stumbling upon this important idea; EPS is almost non-existent in this literature. Also, a fancy dataset will not help. All you need is to recognize the dissonance between how our theories postulate people think and behave and what people in the real world actually think and behave. And access to Compustat, of course.

Once you have some solid research projects going, it is easy to find new research ideas related to your current agenda. In the case of EPS vs NPV, Alex Chinco and I noticed that managers are not the only ones caring about EPS. When you read analyst reports (yes, actually download and read the reports), most of their discussion revolves around EPS. Again, we spotted a discrepancy between how economists think analysts calculate price targets and how analysts perform this task in reality. To the surprise of many economists,

analysts do not discount cash flows; rather, they multiply the EPS they predicted by a trailing (i.e., historical) price-to-earnings ratio (Ben-David and Chinco, 2024).

3 Systematic Approach to Idea Generation

3.1 Casting a Wide Net: Embracing Random Inspiration

Impactful ideas emerge when preconceived priors meet *conflicting* anecdotal evidence from the real world. When such conflict exists, there are grounds to ask what in our priors is wrong and whether there is a better way to think about the world. This is how new theories come to existence and replace old ones; Think, for example, how behavioral economics offers some better predictions about household behavior than neoclassical economics does.

These anecdotes could come from diverse and often random sources. For instance, a conversation with a friend at the dog park, a trader of energy futures, led me to ponder: What determines prices in this market? Is it efficient? What role do arbitrageurs play? Is a market for energy contracts beneficial or detrimental to end consumers?

To discover new ideas in a systematic fashion, I am avid user of an RSS reader, a tool for discovering new content. RSS apps push updates from a curated pool of blogs and newsletters directly to your device. I use Feedly, but there are excellent free RSS readers like Inoreader, Vienna, NetNewsWire, Lire, and NateParrott, some with AI capabilities. Start by adding blogs from a Google search such as "best 100 economics blogs" and expand your list over time.

I seek out blogs and newsletters that analyze current issues and explore novel topics. I follow sources from sociology, psychology, and beyond, including varied political and ideological perspectives. I prefer sources that publish weekly or monthly to avoid information overload. My aim is to cast a wide net, like a fisherman, to capture diverse insights. It will be sufficient to spend 20 minutes a day scrolling through your RSS reader. While I don't read every blog post, I often find articles that offer valuable real-world insights, contrasting

with established economic theories.

Sifting through blogs, white papers, and newsletters (rather than *The Journal of Finance*) brings fresh perspectives. I often find new insights in areas I didn't anticipate. The next intriguing fact might relate to interest rates, the cost of higher education, or 21st century leisure trends. Start using your RSS reader, and I guarantee you will be surprised by how differently the world operates compared to what traditional economic theories predict.

3.2 Rapid Evaluation: Discard Ideas Efficiently

The second pressing question is how to shorten the research cycle. Having a flow of interesting ideas would be useless if assessing each idea's viability takes months. Typically, a PhD student surveys the literature, collects data, cleans the data, runs regressions, writes a summary, and then presents it to their advisor. This process is very lengthy and frequently yields lukewarm or disappointing reactions.

Students often view research projects as composed of two components. The first is an idea expressed as an *interesting relationship* between two objects, as in "the relation between bitcoin and inflation seems very interesting." This is actually not a research idea, but rather an illusion of one. A "relation" between two objects is not a research question. The second component is a novel dataset that could be informative about the relationship between those two objects. As discussed above, a new dataset is not likely to result in a research idea (although it is much fun to work on data!). Tinkering the data is the most time demanding part of research project, and should be deferred as much as possible.

If indeed we can generate a flow of potential research ideas, we need a quick and efficient method to screen them. In other words, we need a process that allows us to quickly discard ideas that initially seem interesting but are unlikely to become solid research projects.

I like thinking about the goal of this process in terms of Type I errors (false positives) and Type II errors (false negatives). For a creative researcher with many ideas, Type I errors are costly, as time spent on a poor project could have been spent on more promising

ideas. Type II errors are less costly—you might occasionally miss on a good idea, but many other interesting ideas will eventually show up in your mental pipeline. Therefore, when evaluating new ideas, I maintain a very high bar, especially ahead of project stages that require significant time investment.

When a new idea arises, I imagine the best possible paper I could write and write down a short description of it. The simplest way is to write an imaginary abstract, sometimes called a précis. This process involves spelling out the motivation, the data to be used, the "ideal" experiment, the imagined results, and the takeaways.

Writing down research ideas forces you to explain why they are interesting and how you plan to test them. Importantly, the goal isn't to reverse-engineer the envisioned results, but rather to filter out ideas that initially seem reasonable in the abstract but collapse when spelled out coherently. If you can't turn your idea into at least one exciting imaginary abstract, there's no point in starting the project.

Once you have an abstract, you need to evaluate whether it is engaging, intriguing, and novel. This might be the right time to seek feedback from colleagues, advisors, or coauthors.

Assuming that you feel confident about your abstract (and hopefully your advisor does too), you need to look for encouraging signs in the data that your project is viable. When I say "data," I don't necessarily mean a thousand lines of code. What you need is some assurance that you are on the right track. It can be, for example, an aggregate statistic that you recover from a white paper that will increase your confidence in your hypothesis and the viability of the paper.

At this point, you should also start thinking about the necessary conditions for your project to succeed. For example, if your project relies on a natural experiment (as in the Wells Fargo example below), you should find supporting evidence that the first stage can be detected in the data.

Research projects should progress incrementally, where at each stage you assess the path forward and potential pitfalls. This approach mirrors the real options method used in drug

development: starting with lab experiments, advancing to animal trials if successful, then to small-scale trials, and finally full-scale trials. As the confidence in success increases, so does the investment in the project. There is no point in going all in only to discover that a critical component (e.g., the first stage in a natural experiment) is missing.

3.3 Example: Wells Fargo Fake Accounts Scandal

In 2016, I came across blog posts discussing the "Wells Fargo fake accounts scandal." Wells Fargo employees had allegedly created millions of unauthorized accounts to meet sales targets and earn bonuses. This event sparked questions in my mind that I felt the literature hadn't adequately answered. I imagined myself receiving an unsolicited credit card from Wells Fargo. Would this improve or harm my financial situation? There were arguments on both sides: increased credit could help fund a business venture, but it might also tempt me into reckless spending.

This thought experiment highlighted a broader question: Does additional credit enhance consumer welfare? The answer likely depends on individual household characteristics, which seemed underexplored in the literature. So, I delved deeper. Some studies suggested that many households are credit constrained, implying that extra credit could be beneficial. Other papers showed that increased credit often leads to wasteful consumption. I noticed that previous research had not used "exogenous" shocks in credit availability. This gap suggested that the literature might not accurately reflect the real world's complexities around credit constraints.

I realized that the Wells Fargo scandal could be an opportunity to study this issue. The first step was to write an abstract outlining why the study would be worth reading, how it could bridge the gap between existing knowledge and reality, and its broader implications. This process often involves visualizing an "ideal experiment" and thinking through potential outcomes, which, in turn, helps identify additional implications.

Here's the imaginary abstract I drafted:

- **Motivation/Tension:** A common belief is that credit constraints prevent households from meeting their consumption and investment goals. Thus, relaxing these constraints should, if households are rational, improve their welfare (they can always refuse the additional credit).

- **Empirical Setting and Data:** Using Wells Fargo's aggressive cross-selling strategy between 20XX and 20XX, we match WF clients with similar non-WF clients and compare their financial and consumption outcomes in the years following. Treated households received unsolicited credit cards, allowing them access to more credit than they had before.

- **Results:** We find that WF clients experienced an abnormal XX% increase in available credit due to the bank's strategy. Treated households were XX% more likely to purchase new vehicles shortly after. However, after 12 months, XX% had accumulated significant credit card debt, averaging XX% of their new credit. Additionally, treated households were XX% more likely to be late on payments, and XX% defaulted on their credit cards. Each dollar of additional credit translated into $XX of annual interest and $XX in late fees. The effects are materially stronger for low-income households.

- **Conclusion:** Overall, an exogenous increase in access to credit caused most households to experience lower welfare.

The goal isn't to draft a flawless abstract but to create a feasible framework for a potentially interesting study. If writing the abstract feels impossible, it's a sign that the idea lacks substance or fails to resolve a key research question.

This mindset guides my exploration through the entire life of a project: constantly assess whether to keep investing or cut losses and move on. I'm not worried about abandoning projects because curiosity constantly fuels new ideas.

Writing down the abstract above also made me realize that the WF natural experiment is not perfect. For example, the relevant WF clients were new clients who opened new bank

accounts. Are they sufficiently similar to existing WF clients and to the control group? I wrote down these concerns, but convinced myself that the issue could potentially be resolved with the right data. In any case, I would need to address this econometric issue down the line.

At that point, I was relatively satisfied with the abstract and decided to push forward. A quick literature search confirmed that no existing work had studied this kind of "natural experiment" involving unsolicited credit cards.

Data exploration came next, a crucial checkpoint. Data analysis can be time-consuming, so it's vital to have confidence in the idea's potential. With my coauthor, we identified a key test for the project: the first stage of the natural experiment. We should have been able to detect a significant uptick in credit card issuance in regions where Wells Fargo had a strong presence. Surprisingly, we didn't find this uptick in the data. To this day, we are not sure why. Perhaps there was too much noise in the data, or these cards were not widely activated by WF clients, or perhaps we didn't do a good job at isolating new bank clients. Without having the first stage at hand, we decided to abandon the project.

What's the takeaway? The goal is to quickly assess and, if needed, discard unpromising ideas—similar to how firms develop new drugs. They start small, evaluate continuously, and determine early on what critical milestones need to be met to justify continued investment. The key is to test conditions early without excessive resource expenditure, especially time. There is only one you.

4   Looking Ahead: The Endless Search for Ideas

As a PhD student in 2005, I was very stressed about my job market paper. None of my ideas seemed like ones that can become a job-market paper. I spent countless hours studying the academic literature and searching for novel datasets no one else had.

I eventually found an idea that became my job market paper, but this idea wasn't an

extension of an ivory tower paper nor didn't it spring out of an exotic dataset. My idea emerged when I picked up my kids one afternoon from their caretaker. At that time, the housing market was the hottest game in town. My kids' babysitter, who was also a part-time real estate agent, told me about an odd trend: homebuyers were paying more than the asking price, with the excess returned at closing; they called it "cash back." Neither of us could make sense of this.

Curious, I dug deeper by talking to other real estate agents and loan officers and reviewing housing ads. This led me to uncover a widespread tactic of mortgage fraud, where buyers misrepresented property values to lenders to avoid down payments (Ben-David, 2011). What seemed like a peculiar anomaly turned out to be a systemic issue.

I didn't intend to write about mortgage fraud and I wasn't particularly seeking ideas about the housing market—instead, the idea found me. But I bet I am not alone. Ask any senior professor about where they source their early research ideas, and you'll find that many have similar stories: seemingly innocuous encounters that unexpectedly sparked significant research projects.

My goal here is to ignite your curiosity as a researcher. Academic research should be seen as an opportunity to learn about the world, not as a chore. There is so much we don't know. To develop impactful research ideas, you need to open your eyes. Unlike quantum particles in physics, the objects we study in economics are all around us, whether in the cafeteria, at the dinner table, or in the dog park. Hopefully, RSS readers can help accelerate the flow of new ideas. We need to be receptive and observant.

I advocate for shortening the research cycle. This approach allows you to explore more ideas and focus on those with the highest potential impact. I can only juggle so many ideas in my brain simultaneously, so project terminations are part of a constructive destruction process: By discarding ideas with little prospect of success, I free up mental space for exciting new ones. When I discard a potential idea, I make room for new ones to emerge.

At the beginning of an academic career, it's natural to worry that once an idea is dropped,

nothing else will come—that you'll face an academic desert. This isn't true. Over time, you'll gain confidence in your ability to find new ideas.

In short, just be curious about the world around you. Your next great research idea is waiting to find you.

References

Ben-David, Itzhak, 2011, Financial constraints and inflated home prices during the real estate boom, *American Economic Journal: Applied Economics* 3, 55–87.

Ben-David, Itzhak, and Alex Chinco, 2023, Modeling managers as EPS maximizers, Working paper, The Ohio State University.

Ben-David, Itzhak, and Alex Chinco, 2024, Expected EPS × Trailing PE, Working paper, The Ohio State University.

Caballero, Ricardo J, 2010, Macroeconomics after the crisis: Time to deal with the pretense-of-knowledge syndrome, *Journal of Economic Perspectives* 24, 85–102.

Graham, John R, 2022, Presidential address: Corporate finance and reality, *Journal of Finance* 77, 1975–2049.

Lowry, Michelle, 2024, The questions being asked: Academic research, the media, and regulators, *Financial Review* 59, 549–560.

Weisbach, Michael S, 2021, *The economist's craft: An introduction to research, publishing, and professional development* (Princeton University Press).