RESEARCH-DRIVEN INSIGHTS ON BUSINESS, POLICY, AND MARKETS

Chicago Booth Review

FALL 2024

Ten ways investors are, or should be, using large language models

Plus: The case for a retro tax code, and an argument for less debate

“Most commentary is pretty awful.”

Page 66

MAKE AI YOUR FINANCIAL DETECTIVE

Financial regulatory filings are rarely scintillating reading. Replete with pro forma language, they often convey little except the punctiliousness of lawyers and compliance experts. Analysts, investors, and regulators wade their way through these reports and do their best to glean valuable information.

But large language models, such as those powering ChatGPT, can "read" these filings by the thousands, which is why LLMs are fast becoming an indispensable investing tool. Research highlights how LLMs can pinpoint what's new and important in a filing, write useful summaries of drawn-out disclosures, and make sense of the words and numbers.

Chicago Booth Review regularly publishes articles and videos about how artificial intelligence is transforming the economy, and in this issue, we've focused on the flood of recent research that is specifically analyzing the market-relevant abilities of LLMs. Our cover story (page 28) highlights 10 projects of note involving Chicago Booth researchers. It's essential reading for anyone looking for an edge or who wants to understand what is driving financial markets.

If technology is one seismic issue facing our economy, politics is another. As this issue goes to press, we are headed into a US presidential election. Whatever the result, the election will likely be followed by a fight over taxes. The Tax Cuts and Jobs Act of 2017, colloquially known as the Trump tax cuts, included a number of provisions set to expire in 2025, which makes a national tax debate all but guaranteed.

Could the United States move forward by looking to the past? In a feature (page 40), we examine a proposal from Princeton's Owen Zidar and Booth's Eric Zwick that involves Congress embracing elements of the tax code circa 1997–policies the researchers contend could be palatable to both the political Right and Left.

As well as bringing you the latest research findings, we're also dedicating more space to new books authored by Booth faculty. Two are excerpted in this issue. In one, Booth's Raghuram G. Rajan, who was the 23rd governor of the Reserve Bank of India, and Cornell's Rohit Lamba outline a path for India's future economic development (page 54). In the other, Booth's Matthew Notowidigdo and Boston University's Tal Gross dive into the economic conundrum at the heart of US healthcare costs (page 60).

Our print magazine is only one way to find out about the latest academic research and thinking. To keep up between issues, visit our website, sign up for our weekly email newsletter, and be sure to follow our social media channels.

Hal Weitzman
Executive director, Intellectual Capital
Editor-in-chief, *Chicago Booth Review*
hal.weitzman@chicagobooth.edu

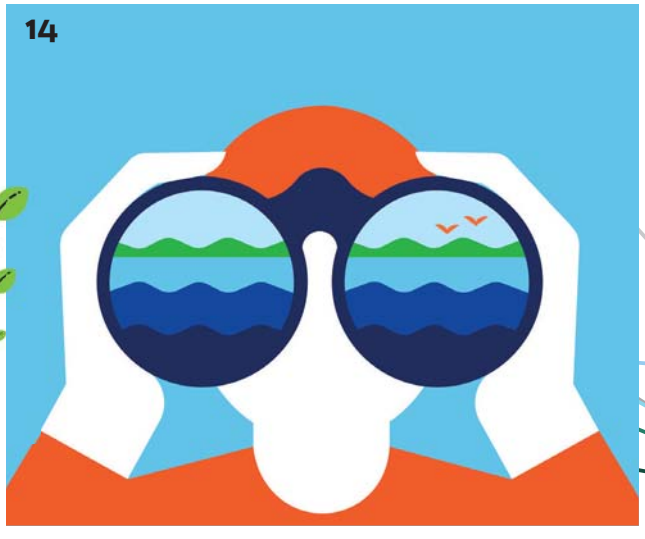
Emily Lambert
Director, Intellectual Capital
Editor, *Chicago Booth Review*
emily.lambert@chicagobooth.edu

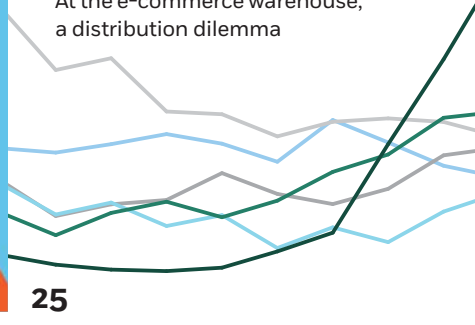
CHICAGO BOOTH
THE UNIVERSITY OF CHICAGO BOOTH SCHOOL OF BUSINESS
chicagobooth.edu/review

COVER ILLUSTRATION BY MADE UP STUDIO

DEPARTMENTS

DATAPOINTS





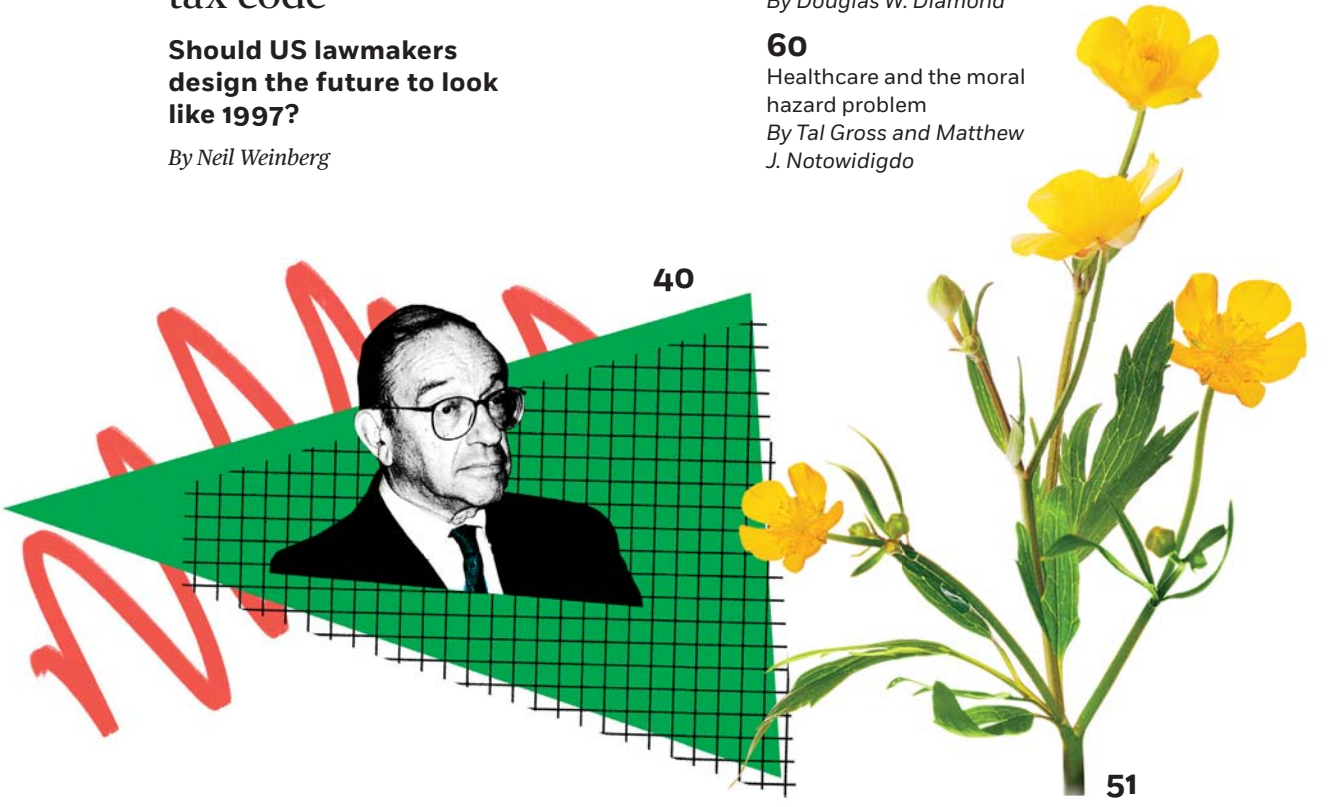
Featured Faculty

Rad Niazadeh, assistant professor of operations management and an Asness Junior Faculty Fellow, studies the interplay between algorithms, data, and incentives in real-time operational scenarios pertaining to platforms and nonprofit organizations. Previously a visiting researcher at Google Research, and currently a faculty member of the Toyota Technological Institute at Chicago (by courtesy appointment), he has won several awards for his research. Video lectures of his PhD courses are available online. (Page 7)

Valeri Nikolaev, the James H. Lorie Professor of Accounting and an FMC Faculty Scholar, is interested in the potential of artificial intelligence to assist investors in making well-informed decisions, thereby fostering more efficient and equitable capital markets. A senior editor of the *Journal of Accounting Research*, he joined the Booth faculty in 2007. (Page 28)

FEATURES

FOOTNOTES



FACULTY ILLUSTRATIONS BY JOHN KENZIE

Eric Zwick, professor of economics and finance, explores the interaction between public policy and corporate behavior, with a focus on fiscal stimulus, taxation, and housing policy. Some of his past research, featured in *CBR*, has detailed the source of wealth for the richest Americans and the importance of pass-through businesses, both of which inform his recent findings about the Tax Cuts and Jobs Act of 2017. (Page 40)

Jane L. Risen, the H. G. B. Alexander Professor of Behavioral Science and a John E. Jeuck Faculty Fellow, conducts research in the areas of judgment and decision-making, intuitive-belief formation, magical thinking, stereotyping and prejudice, and connecting across lines of difference. Her essay about the need for more dialogue is drawn from her commencement address to this year's graduates of the Evening, Weekend, and Executive MBA Programs. (Page 51)

Find the articles to which these comments refer at **chicagobooth.edu/review**.

WE WELCOME LETTERS

Send email to **review@chicagobooth.edu** or send letters addressed to *Chicago Booth Review* at any of the following addresses:

5807 S. Woodlawn Ave.
Chicago, IL 60637

One Bartholomew Close
Barts Square
London EC1A 7BL
United Kingdom

168 Victoria Road
Mount Davis
Hong Kong

Comments may be edited for clarity and/or space.

JOIN THE CONVERSATION

Subscribe to our weekly email newsletter at **chicagobooth.edu/review**, and follow us on social media:

X
@chicagoboothrev

Facebook
@chicagoboothrev

Instagram
@chicagoboothreview

LinkedIn
Chicago Booth Review

YouTube
Chicago Booth Review

Threads
@chicagoboothreview

Magazine: "Inflation affects behavior decades after prices have stabilized" (Summer 2024)

My grandma lived through the Great Depression. It affected her savings patterns for the rest of her life. I spent a lot of time with her as a kid and I also save, but not as much as she did. So I'd believe that family members would pass down the lessons of hyperinflation as well.

—Brooke L. Allen

I hope it does. I hope people learn what my grandparents taught me after seeing the Depression, World War II, and the steel collapse: Pay off your debts. Stay small. Keep a pantry. Stay on good terms with the neighbors. Be able to do stuff. Be good enough you can get help if you need it, but never depend on anyone.

—Jessica O'Leary

A CBR performance review

The content and method of presentation for *CBR* are so good, I keep a hard copy in the trunk of my car to read during the seemingly endless time I spend waiting for medical appointments here in the Valley of the Sun.

The brief but complete synopses remind me of my Booth Executive MBA sessions with the likes of Professors Bud Fackler and Merton Miller, both of whom became my friends. They created an economic panorama through which to filter all the chatter in the present info explosion.

Keep up the brilliant research. I will continue to read via the email newsletter and my hard copy.

—John M. Gleason Jr.

Magazine: "Some students do better in remote learning" (Summer 2024)

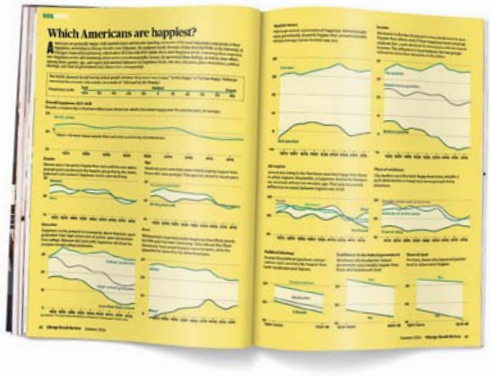
I wish this were discussed more. I am a special-education teacher. Most of my students have learning disabilities or other health impairments affecting education. Many of them flourished in an online learning environment, especially if they didn't have to keep their cameras on. Many also flourished with asynchronous learning because they could complete things on their own schedule.

—Amanda Manda

Why are we still only measuring learning with test scores?

—Cascade DuSel

Magazine: "What rice farming can teach us about happiness" and "Which Americans are happiest?" (Summer 2024)



In the Summer 2024 issue of *Chicago Booth Review*, there are two articles on happiness. In neither is there any indication of what the word *happiness* means, much less the reliability of how it is measured.

Does *happiness* mean the extent to which one is satisfied with one's life, the frequency and intensity of experiencing joy or pain, or something else? The usual method for measuring happiness is to ask respondents how happy they are. Is that a reliable measure? Are individuals really able to determine how happy they are compared with others?

What is the value of articles ranking undefined concepts without reliable methods of measurement? Does possible reader interest justify carrying them?

—Milt Lauenstein

NOW ACCEPTING APPLICATIONS FOR OUR TWO SPECIALIZED MASTERS PROGRAMS

Now is the time to apply to our Master in Management and Master in Finance Programs. These Specialized Masters Programs are for recent college graduates with limited work experience, allowing individuals to begin their careers with confidence. Both are taught by our world-class faculty, both are hosted in Chicago, and both allow students to tap into our global alumni network.

MASTER IN MANAGEMENT

The MiM Program is a 10-month graduate program designed for high-achieving new graduates in the liberal arts or STEM fields. Taught by the world's leading business school faculty, the program provides students with core business skills that they can combine with their undergraduate experience to enhance their competitiveness as they enter the job market.

Learn more about the Master in Management

MASTER IN FINANCE

The MiF Program is for students with quantitative backgrounds who are searching for a masters degree that builds on their analytical aptitude and allows them to start a career in finance with a competitive advantage over their peers. Students learn from the world's top finance scholars and complete the program in 15 months.

Learn more about the Master in Finance

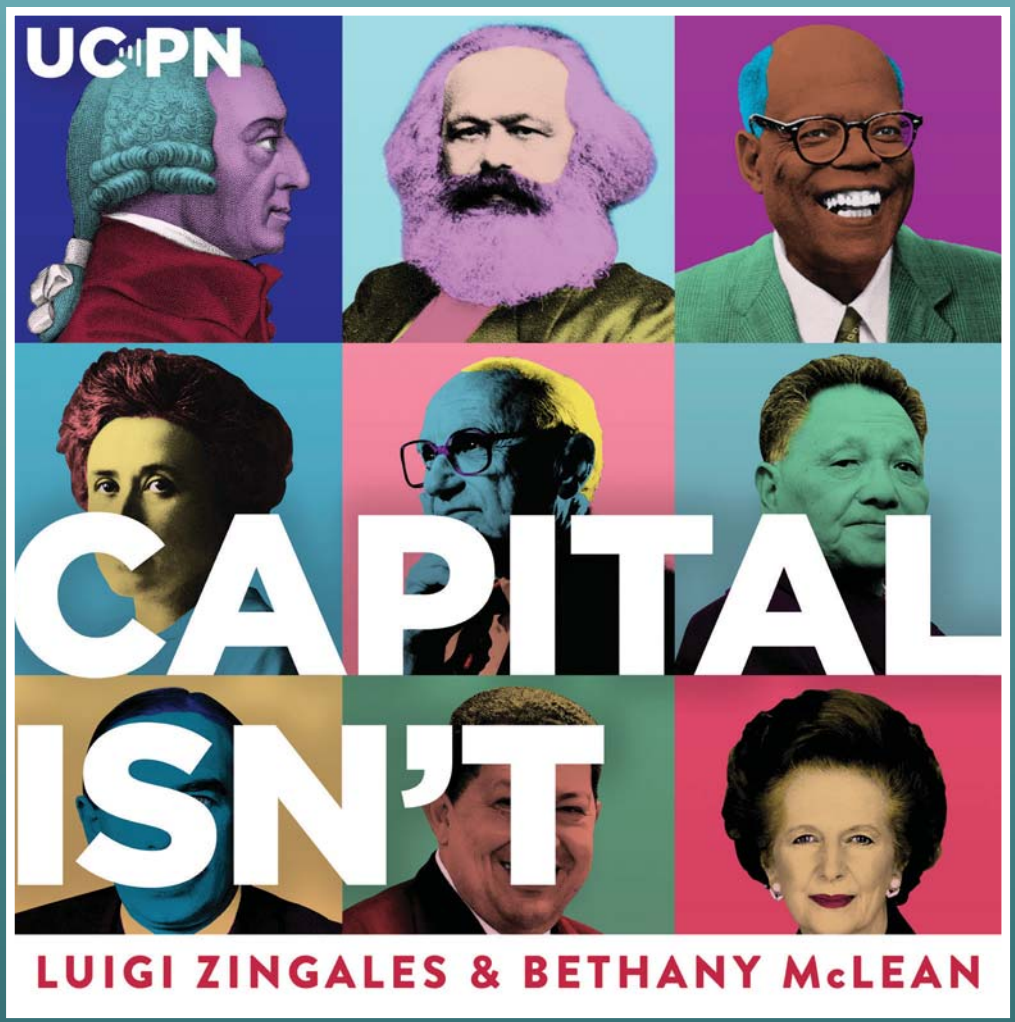
DO YOU KNOW COLLEGE SENIORS OR RECENT GRADUATES WHO ARE LOOKING TO KICK-START THEIR CAREER?

Recommend promising candidates:
chicagobooth.edu/shapetheclass-s25

Learn how your organization can partner with Booth students and/or hire alumni:
chicagobooth.edu/contact-corporate-relations

Questions? Email us at **masters@chicagobooth.edu**.

CHICAGO BOOTH
THE UNIVERSITY OF CHICAGO BOOTH SCHOOL OF BUSINESS



IS CAPITALISM THE ENGINE OF DESTRUCTION OR THE ENGINE OF PROSPERITY?

Hosts Luigi Zingales, a world-renowned economics professor, and Bethany McLean, a *Vanity Fair* contributing editor, explain how capitalism can go wrong, and what we can do to fix it.

Listen and subscribe wherever you get podcasts, or on **capitalisnt.com**.

DATAPOINTS

Which workers will benefit from AI?
Page 13

Charge drivers to improve public transit
Page 22

Algorithms and AI can make hiring more diverse

The cost is likely minimal to achieve a fairer outcome

Many companies are searching for tools to help them hire diverse, productive workforces. Even if diversity is not the main hiring goal, they may want to ensure they're not overlooking talented individuals because of systemic discrimination, says Chicago Booth's Rad Niazadeh.

Automated, data-driven algorithms, in conjunction with AI and machine learning, can support organizations in these efforts, suggests research he conducted with Booth PhD student Mohammad Reza Aminian and Yale's Vahideh Manshadi. The increasing use of algorithms in hiring has raised concerns that they may reinforce human biases due to implicit bias present in the data they use. (For more, read "AI is only human," in the Summer 2019 issue and online at **chicagobooth.edu/review**.) However, the researchers demonstrate that algorithms designed with fairness-and-diversity constraints can guide companies to interview a more diverse set of candidates and extend employment offers to a broader range of people—with a minimal cost.

ILLUSTRATION BY CHRIS GASH

Aminian, Manshadi, and Niazadeh propose an algorithmic framework for screening and hiring that includes a number of such constraints. Their framework is for sequential processes, meaning those in which candidates are evaluated one after another rather than all at once.

The researchers started by analyzing well-established "candidate priority indices"—also known as Weitzman indices, thanks to the pioneering work of the late Martin L. Weitzman in 1979. According to classical economics, hiring managers should use the Weitzman indices to devise an optimal strategy for interviewing and hiring candidates. Through an in-depth theoretical analysis, Aminian, Manshadi, and Niazadeh find that to make hiring outcomes fair and diverse, managers need to adjust these indices by increasing the priority of candidates from disadvantaged populations in a specific way and modulating the priority of other candidates.

An organization's exact goals will drive the specific constraints and adjustments, the researchers argue. For example, if an engineering company wants to hire more high-quality female candidates, it's not enough for it to simply interview more women. A more refined diversity constraint would prompt the company to include more high-quality female candidates in its interview pool. This would avoid tokenism, which is a way of gaming systems that require diversity and inclusion.

To demonstrate the applicability of their framework beyond theory, the researchers ran simulations of the algorithm, imposing various fairness and diversity constraints. In these simulations, hypothetical job candidates were marked as members of either a disadvantaged or a privileged demographic group.

The researchers estimated the candidates' "quality" by assigning each person a short-term and a long-term score. The short-term score reflected formal qualifications (such as educational background). These scores were unequally distributed across all the candidates to reflect the impact of privilege on access to high-quality education and other resources. In contrast, the long-term score estimated the true quality that a person provides over time, due to characteristics—including intelligence, work ethic, and ambition—that the researchers assumed to be equally distributed across demographic groups.

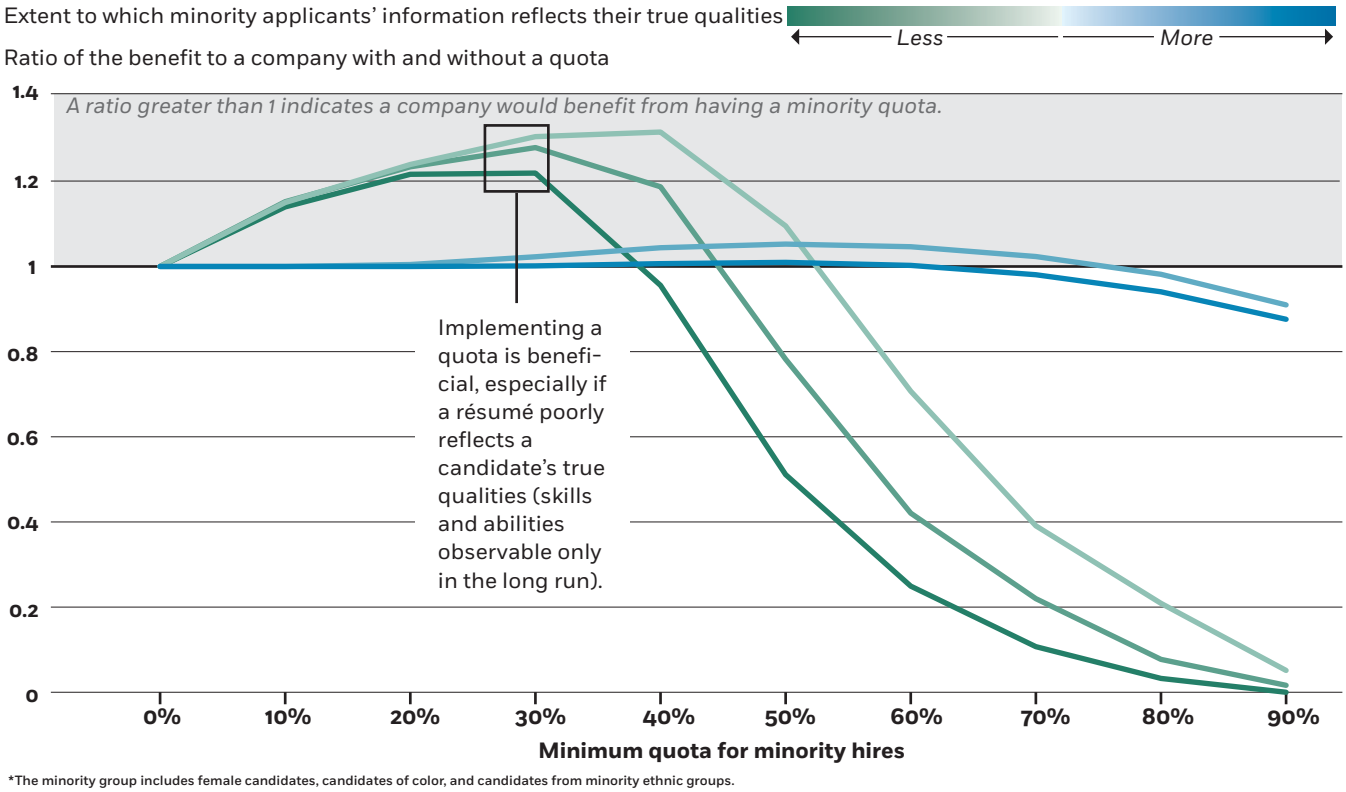
The algorithm could see only the short-term scores, mirroring how real-world recruiters assess candidates on the basis of their résumés or interviews. But the researchers then used each candidate's true-quality score to measure which of the hiring practices yielded more benefit in the long run. Which approach, by these simulations, truly led organizations to hire candidates who had the highest long-term scores?

The findings suggest that automated, data-driven algorithms incorporating

An upside for quotas

The research indicates that by countering biases, quotas for minority candidates can lead companies to hire people who would benefit the organization in the long run and might otherwise be overlooked.

Long-term effect of a minority quota*



*The minority group includes female candidates, candidates of color, and candidates from minority ethnic groups.

Aminian et al., 2024

fairness and diversity constraints can lead companies to hire people who appear, on paper, to be less qualified than candidates brought on through a process that ignores demographics.

But even in terms of employee quality on paper, the cost to a company is likely minimal to achieve a fairer outcome, according to the research.

"If you force a company to hire, on average, 10 women for every 10 men, you might reduce the number of top candidates they hire, such as those with the highest GPA or a degree from an Ivy League school, simply because you added an extra constraint to the search," says Niazadeh. "But, in reality, you might not hurt the utility of the search by much."

He explains that there may be several optimal ways of hiring people, and while a demographics-blind policy yields the best results in terms of short-term scores, other methods are still reasonable.

The simulations also suggest that this kind of inclusive practice benefits organizations in the longer run: imposing quotas, even when one group boasts stronger qualifications than another, produces a better workforce (as measured by the hypothetical candidates' true quality) than hiring on the basis of short-term scores alone. An organization will find better employees if it recruits, say, a 50/50 male-female team in which 16 out of 20 boast Ivy League degrees than if it hires 20 Ivy League graduates, the majority of whom are men.

"Imposing socially aware constraints such as demographic parity or [a] quota can even make the search *more efficient* in terms of true unobserved qualities," write the researchers.

The exception comes when extreme constraints are imposed in settings where systemic discrimination has created vastly disparate groups in terms of formal qualifications—for example, if the demand were that 10 Black STEM PhDs be hired for every 10 white STEM PhDs, despite the fact that, according to a report commissioned by the Alfred P. Sloan Foundation, only 5 percent of PhD holders in the science, technology, engineering, and math fields were Black as of 2021. Under circumstances such as these, the simulations reveal, positions often go unfilled, reducing the long-term utility of a team because the team itself is smaller than it should be.

Many people have thought about algorithmic fairness in decision-making, says Niazadeh. "When it comes to designing machine-learning algorithms for high-stakes applications such as loan decisions, computer scientists and economists have studied algorithms that favor disadvantaged groups. This is in response to evidence that demographics-blind ML algorithms discriminate due to skewed data," he says. But the "fair" ML algorithms have tended to make straightforward choices based on one-time signals—for example, deciding whether a loan application gets approved on the basis of a potential borrower's credit history.

Hiring decisions are often more complex in nature. Here, it takes time and resources beyond scanning a résumé to find out if a candidate is any good. Markers of quality are dynamic, since a hiring manager's opinion of each candidate may change after a first interview, a second interview, and a site visit.

"That's the technical challenge," says Niazadeh. "Hiring a person is more complicated than opening Door 1, 2, or 3 and seeing what you get." The researchers argue that the complexity calls for a Markovian scheduling framework. (A Markovian model, named for its creator, the late Andrey Markov, describes a sequence of events in which the probability of the next depends on the outcome of the previous one.) This framework goes beyond static ML problems and even Weitzman's indices.

While the researchers' algorithmic approach has the potential to influence hiring in many countries, especially when it involves a sequential search process such as in executive recruiting, Niazadeh predicts that US organizations might balk, given the political and legal questions around diversity and inclusion. Even those open to quotas may find the inner workings of the tools uncomfortable, he adds, because they rely on a degree of randomness: when two candidates appear to be equally qualified, the algorithm essentially flips a coin.

But he says that some policymakers have agreed to use randomization in selecting citizens for assemblies or juries and in distributing legislative seats. This approach, he says, helps achieve the optimum outcome under the fairest conditions, on average.—*Rose Jacobs*

Go to **chicagobooth.edu/review** to see citations for research mentioned in this article.

BIG FOOD WANTS YOU TO OVEREAT

"Food companies are not social-service or public-health agencies. They're businesses with shareholders to please. Wall Street has adopted the shareholder-value movement, so the primary goal of corporations is to sell products and make profits.

The only thing that matters to food corporations is selling as much food as possible at as high a price as they can get away with. To do that, they created an environment that encourages people to eat too much. They made it socially acceptable to eat everywhere, anytime, and in large amounts. Portions alone are a sufficient explanation for obesity. You don't need anything else to explain it."

—MARION NESTLE, of New York University, in edited remarks from an episode of the *Capitalisn't* podcast, presented by Chicago Booth's Stigler Center for the Study of the Economy and the State

Do diverse leadership teams produce better performance?

When companies make hiring and other decisions related to diversity, many cite the positive impact they expect these choices to have on the business's performance. A 2022 analysis of Fortune 500 companies' websites by Boston University's Oriane Georgeac and London Business School's Aneeta Rattan finds that about 80 percent use diversity's impact on performance as their main justification.

But does that argument actually hold up? University of North Carolina's Sekou Bermiss, Texas A&M's Jeremiah Green, and UNC's John Hand (a visiting professor at Chicago Booth since 2017) analyzed data for the full set of companies in the S&P 500, and they find no evidence of a relationship between greater diversity on executive teams and better subsequent financial performance.

Consultants, business leaders, and activists often promote what the researchers call the "business case for diversity," or the notion that greater racial and ethnic employee diversity yields financial benefits for the employer. Management consultancy McKinsey & Company, for example, has several studies—released in 2015, 2018, 2020, and 2023—that report a large and statistically significant positive relationship between adjusted earnings at an anonymized set of large public companies and the diversity of their executives.

"What the data shows is that companies that have more diverse leadership teams are more successful. And so the leading companies in our data sets are pursuing diversity because it's a business imperative and driving real business results," Dame Vivian Hunt, formerly a senior partner at McKinsey, said in a 2018 interview on *Bloomberg Surveillance*.

But such statements contrast with more nuanced findings in academic research about the costs, benefits, risks, and returns of greater diversity, write Bermiss, Green, and Hand, who decided to do their own investigation.

The researchers gathered information from corporate websites about the racial and ethnic makeup of S&P 500 leadership teams, as recorded midyear in 2011, 2014, 2017, 2021, and 2022. They then determined if any of nine measures of executive racial and ethnic diversity (including McKinsey's) predicted variation in any of six measures of financial performance—sales growth, gross margin, profit margin, return on assets, return on equity, and total shareholder return—over the subsequent fiscal year.

They also included the fraction of corporate executives who are women in order to simultaneously test another commonly voiced claim: that greater gender diversity on executive teams leads to better financial performance.

Their analysis spanned the period that followed the 2020 murder of George Floyd, when many companies made a public commitment to racial diversity. Yet their research finds no positive relationship between either racial and ethnic diversity or gender diversity and financial performance over the next fiscal year.

No clear link

The researchers looked at how nine measures of leadership diversity affected performance outcomes at S&P 500 companies over five years, running 270 regressions in all. A diversity measure significantly predicted performance in only nine instances, they find.

Statistical significance of diversity on financial performance

Significant predictor of performance in one of the five years: ● Positive ● Negative ● Not a significant predictor in any year

Diversity measures	Company performance measures: Sales growth	Gross margin	Profit margin	Return on assets	Return on equity	Shareholder return
McKinsey's measure of executive concentration across eight racial and ethnic categories	●	●	●	●	●	●
McKinsey's measure of executive concentration across five racial and ethnic categories	●	●	●	●	●	●
Proportion of all non-white executives	●	●	●	●	●	●
Shannon entropy measure of executive-distribution evenness across racial and ethnic groups	●	●	●	●	●	●
Measure comparing diversity of executives with diversity of graduates from top-tier universities	●	●	●	●	●	●
Proportion of Black executives	●	●	●	●	●	●
Proportion of East Asian executives	●	●	●	●	●	●
Proportion of Latino executives	●	●	●	●	●	●
Proportion of South Asian executives	●	●	●	●	●	●

Bermiss et al., 2023

The researchers ran 270 regressions in all (the nine racial/ethnic leadership diversity measures applied to six performance measures for each of the five fiscal years studied). Diversity predicted the next year's financial performance in just under 5 percent of the regressions. "The data speak almost exactly to what would be expected to be seen in terms of results purely by chance," says Hand. Analysis of the gender diversity in executives produced similar results.

As such, their findings do not support the business case for either racial and ethnic or gender diversity, the researchers argue. McKinsey's results, they contend, indicate that higher financial performance leads to greater diversity among executives, not the other way around. "McKinsey found that those in their particular set of firms that were more profitable made the time to become more diverse," adds Green. "The fewer existential crises a firm is dealing with, the more time and money it can spend on matters of diversity."

Green and Hand, in a separate project, tried but failed to replicate McKinsey's results for the S&P 500.

Reached for comment, a McKinsey spokesperson issued the following statement:

> *In light of a recent study criticizing our methodologies, we have reviewed our research and continue to stand by its findings–that diversity and inclusion are associated with a higher likelihood of financial outperformance. We have also been clear and consistent that our research identifies correlation, not causation, and that those two things are not the same. For more than a decade, we have published groundbreaking research into the business and economic impact of a diverse workforce, and welcome discussion of this important topic.*

Bermiss emphasizes that the research he conducted with Green and Hand shouldn't be interpreted as saying that companies ought not to hire diversely. "There are benefits to diversity in the workplace that are not directly related to firm-level profitability," he says. "However, firms should be clear about the motivations to implement diversity policies." Perhaps relying solely on the business case for diversity is a shaky justification for hiring–after all, he asks, if a company says it's hiring diversely because doing so improves profit margins, what happens if those profit margins do not improve? "It would likely reduce support for these policies," he concludes.–*Rose Jacobs*

Go to **chicagobooth.edu/review** to see citations for research mentioned in this article.

Job seekers valued an employer's emphasis on socially conscious policies about as much as they did a pension.

ENVIRONMENTALLY CONSCIOUS COMPANIES ATTRACT SKILLED WORKERS

SOME WORKERS prefer organizations with a commitment to environmental, social, and governance concerns, according to a team of researchers that includes Chicago Booth's Emanuele Colonnelli and Thomas Rauter. While this tends to increase overall wages and economic output, it also further widens income disparities, they find.

Analyzing data from Brazil, the researchers sought to understand the impact of ESG practices on the allocation of talent in the labor market, a key component in determining both a country's economic output and its division of wealth.

They assessed how companies and job seekers view ESG policies. For companies, they relied on a market-research group that gathered responses in July 2023 from nearly 1,100 Brazilian enterprises regarding ESG initiatives and the motivations behind them. Eighty-one percent reported having some form of ESG practices in place, while 41 percent indicated they were "extensively" implementing them.

Meanwhile, through Brazil's leading job-matching platform, the researchers asked prospective employees how ESG policies influenced their preferences. They received about 1,200 responses from users who were demographically similar to the country's workforce as a whole.

The job seekers rated their interest in 20 hypothetical postings on the basis of wages, ESG practices, and other factors. While respondents were aware that the postings were fake, they were told the results would be used to match them to real positions, giving them an incentive to reply truthfully.

Companies didn't necessarily pursue ESG initiatives with the main goal of attracting and retaining talent, the research indicates. However, respondents had a statistically significant preference for companies whose job postings reflected an interest in advancing ESG goals, the researchers find. This signal increased the attractiveness of a job by the same amount, on average, as a 10 percent rise in monthly wages. The job seekers valued an employer's emphasis on socially-conscious policies about as much as they did a pension or food allowance and 60 percent as much as being permitted to work remotely.

But ESG preferences were unevenly spread throughout the workforce. The initiatives were effective in attracting white, highly educated individuals who identified as politically liberal or moderate. These workers expressed a strong interest in environmental issues, but not in governance. The other end of the labor spectrum yielded starkly different results: less-educated, non-white, and politically conservative job seekers expressed no additional interest in working for companies that emphasized ESG factors for any reason.

The implications of these findings go beyond the question of who works where. The sorting of labor due to ESG practices increased both total wages and output, but because skilled workers were found to reap an outsize portion of these gains, it exacerbated economic disparity, the study finds. The researchers estimate that these practices increased the wage differential (the wage difference between employees performing the same job) up to 4 percent relative to an economy without them.—*Neil Weinberg*

Emanuele Colonnelli, Timothy McQuade, Gabriel Ramos, Thomas Rauter, and Olivia Xiong, "Polarizing Corporations: Does Talent Flow to 'Good' Firms?" Working paper, November 2023. Go to **chicagobooth.edu/review** to read a longer version of this article.

Consumers say they care about ESG, but don't spend like they do

Regulators increasingly require companies to report environmental, social, and governance information. Such disclosures provide more transparency for investors, as well as for consumers who want to align their spending with their values. But does knowing about the ESG activities of a business actually change what shoppers buy?

Sinja Leonelli of New York University, Chicago Booth's Maximilian Muhn and Thomas Rauter, and NYU's Gurpal Sran put consumer attitudes to the test. In an experimental survey, they find that companies' ESG disclosures had little impact on customer spending.

The researchers worked with Numerator, a US market-research company that administers consumer surveys. Numerator tracks respondents' purchases in stores and online, making it possible to match survey results with spending habits.

Leonelli, Muhn, Rauter, and Sran collected information from more than 24,000 survey participants, initially asking what factors most influence their purchase decisions. "By far, the two most important purchase considerations are product quality and price," the researchers write. While survey participants said they also consider ESG issues such as carbon footprint, these factors ranked much lower.

The researchers also asked participants directly whether they prefer to buy from "ESG-responsible" businesses. Survey respondents had a moderate preference for purchasing from such companies, but 35 percent of participants said they didn't have information about brands' ESG activities.

This prompted the researchers to put that information right in front of them. They used survey respondents' purchase history from Numerator to build each participant an individualized portfolio of 15 products—some of which the participant had previously bought, and some of which were substitutes for those products.

The information in the profiles varied, with different versions randomly assigned to respondents. Some profiles contained information about the company's ESG performance; some linked to the company's ESG report; and some contained information not related to ESG, such as financial data or product reviews.

For example, the social-oriented ESG profile of the Jimmy Dean Frozen Meat Lovers Breakfast Bowl included information such as: "Tyson Foods offers financial grants and food product donations to regional food banks serving its communities." The environmentally focused profile omitted that fact, but reported on Tyson's use of renewable and recyclable materials. Participants who received profiles with a link to a company's ESG report had to click through to receive any ESG information, and could choose not to click the link.

The researchers asked participants how likely they were to buy each product over the next six months. Those who viewed the full ESG report for a product showed an increase in purchase intent of 0.18 percentage points—the largest increase out of all the types of information received. Participants also reported being more likely to buy a product when they received information about the business's social or environmental activities, and to a lesser extent, its efforts at good corporate governance.

The effects on actual purchases were both small and short-lived, however. In the two weeks after viewing a company's ESG report as part of a product profile, households in the study increased their number of purchases of that product by 1.2 percent. After being told about the social activities of a business, consumers increased purchases by 0.3 percent.

By the third or fourth week after the survey experiment, even those small changes disappeared. To understand why, the researchers sent participants a set of follow-up questions and found that 65 percent of respondents who remembered taking the initial survey said the experiment did not change their behavior, primarily because they did not remember the ESG information or did not have time to consider it.

"There are so many frictions in the real world that prevent this information channel from being as effective as we might wish," Rauter says.

The researchers consider that requiring or encouraging businesses to provide ESG information could indirectly benefit consumers. Investors or other stakeholders might use the information to reward or pressure certain businesses. Regulators also could use disclosures to design certifications or labels that make it easier for consumers to remember and apply ESG information.

Still, "if you want to change something for the good and have firms become more environmentally friendly," Muhn says, "maybe providing more disclosure geared toward consumers is not the right lever."
—Amy Merrick

Sinja Leonelli, Maximilian Muhn, Thomas Rauter, and Gurpal Sran, "How Do Consumers Use Firm Disclosure? Evidence from a Randomized Field Experiment," Working paper, January 2024.

ILLUSTRATION BY MARTIN LEÓN BARRETO

Which workers will benefit from AI?

Economies are always adjusting to transformative technologies—from windmills, steam engines, and electricity to computers, the internet, and now artificial intelligence. Can lessons from those earlier technological revolutions help us anticipate how AI will affect inequality and the labor market?

Chicago Booth's Rodrigo Adão, MIT's Martin Beraja, and University of Texas's Nitya Pandalai-Nayar examined two specific periods of major technological change, the manufacturing revolution of the early 20th century and the transition to computers and the internet in the late 20th century, and find a stark difference in how they played out in the workforce. Their findings may tell us just how pronounced the changes from AI will be.

The researchers used US census data from the late 19th century to 2019, focusing on male workers aged 16-64, to measure the employment and wage response in occupations most exposed to the new technologies, labor-supply adjustments across worker generations, and the ability of current workers to apply the new technologies. They then constructed a model to analyze how the transferability of worker skills affects the economy's adjustment to new technologies.

They find that the manufacturing revolution tapped into existing skills—from agriculture and manual labor—that were easily transferable to factories, making for a rapid transition without widening inequality. Older and younger workers both contributed similarly to the new manufacturing economy; the relative wage in the new manufacturing jobs rose less, meaning less financial inequality on the basis of specific skill sets; and the broad economic benefits appeared relatively quickly.

"If a technology comes in, and it can be used by anyone, it'll be quickly adopted, will have big aggregate effects quickly, positive effects in general, and very little inequality or displacement effects, simply because we all have the skills to use it," Adão says. Slow-adopting technologies may also have big aggregate effects, but they are delayed.

"You could, for example, come from a farm and be doing labor work on the farm, come to a factory and spend a few weeks or months learning how to do the labor work in the factory," he says. "But fundamentally it was using similar skills that people already had that could complement those technologies."

Adjustments to innovations in information and communications technologies, on the other hand, were slower and more unequal, dramatically slowing the transition, the researchers find. That revolution created jobs requiring more cognitive skills and higher education, necessitating the training of a whole new generation of workers and disproportionately benefiting younger people.

The slower the labor-market transition to a new technology, as in the case with ICT, the more the gains will accrue for people in the future, the researchers find. Those most negatively affected will be workers today who may be displaced because they cannot redeploy their skills effectively to participate in the sectors expanding as a result of the new technologies.

"When there is more inequality, it means that the aggregate effects such as GDP growth and overall production are going to take longer to accrue," Adão says.

He suggests that policymakers might consider compensating workers who are losing today through the gains of those who are most likely to win in the future. Such policies could include transfers—or direct payments—to displaced workers that would slow over time as younger workers entered the system, or taxes on new technologies that could fund the social safety net, or programs to help workers learn new skills. (For more on policies to help displaced workers, read "What's the best way to retrain jobless workers?" online at **chicagobooth.edu/review**.)

The researchers don't focus on AI but do note that their findings raise the question of whether the adjustment to the new technology will more closely resemble that of the manufacturing or of the ICT revolution.

The key question, Adão says, is whether AI is more likely to complement skills that already widely exist, making many workers more efficient, or to tap into the skills of only a small part of the labor force. The researchers think it's too early to tell.

What is clear, Adão says, is that if the labor market ultimately adjusts to AI as it did to the manufacturing revolution, it will be a much faster transition. But if it adjusts as it did to the ICT revolution, the transition will be slower, even as the technology's diffusion accelerates.—*Rebecca Stropoli*

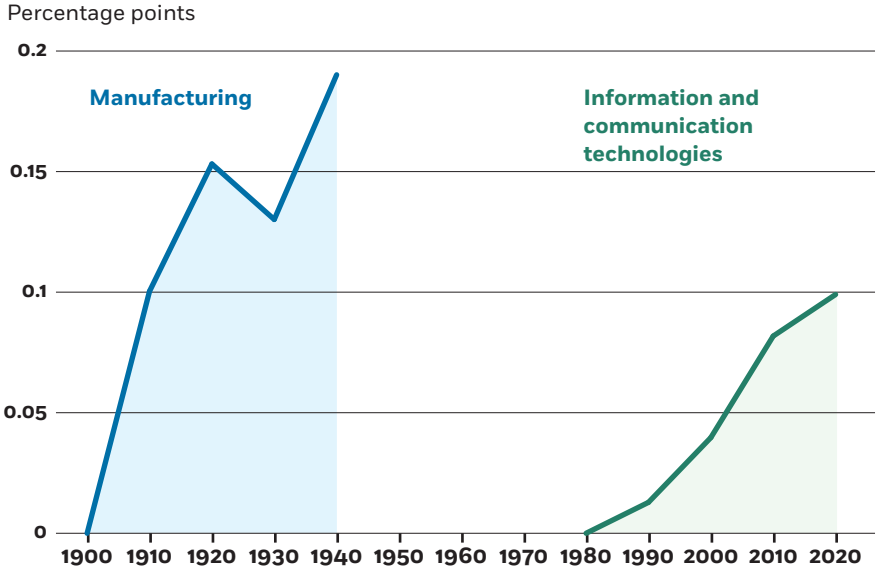
Rodrigo Adão, Martin Beraja, and Nitya Pandalai-Nayar, "Fast and Slow Technological Transitions," *Journal of Political Economy Macroeconomics*, June 2024.

Different reactions in two technological revolutions

Occupations more exposed to new manufacturing technologies in the early 1900s grew rapidly. But jobs that used late-20th-century ICT innovations grew slowly.

Relative employment growth in occupations with 1 standard deviation higher exposure to the new technologies versus other occupations

Compared with the earliest year after the arrival of the new technologies



Adão et al., 2024

WHY YOUR CREDIT REPORT STOPPED SHOWING YOUR CARD PAYMENTS

IF YOU LOOK at your consumer credit report, you may be surprised at what you don't find there. While it will show the amount you paid on your auto loan, mortgage, and unsecured loan, you may not have access to information about how much you paid on your credit card. If so, you are not the only one: 165 million US consumers are missing this information, according to research by Rice University's Benedict Guttman-Kenney and Northeastern University's Andrés Shahidinejad (both recent graduates of Chicago Booth's PhD program).

Between 2009 and 2013, lenders usually shared information with credit bureaus regarding payments—both scheduled and actual—for credit cards, as well as for auto loans, mortgages, and unsecured loans. But by 2015, information on the actual credit-card payments was missing in credit reports for consumers holding cards from any of the six largest lenders, the researchers write. Moreover, the fraction of credit-card accounts in credit reports that showed payments information fell from 89 percent in 2013 to 36 percent at the end of 2022, they find.

What changed? To study this, Guttman-Kenney and Shahidinejad used consumer credit reporting data from the TransUnion Consumer Credit panel, housed at Booth's Kilts Center for Marketing, which is a random sample of anonymized information on 10 percent of the US consumers who have credit reports.

Analyzing monthly credit reporting data from 2009 to 2022, they homed in on Trended Data, a product the credit bureaus launched in 2013.

Trended Data uses a history of credit-card actual payments and statement balances over time to estimate spending (the value of new purchases) and revolving debt (the statement balance less actual payments). Its measures allow lenders to locate and distinguish between two particularly profitable consumers: high-spending cardholders who generate interchange revenue (transaction fees paid by the merchant), and cardholders who produce interest revenue by carrying a high balance without defaulting.

The researchers find evidence that Trended Data led to a breakdown in information sharing. The product made it easier for lenders to poach customers from each other, and especially from the largest and most profitable among them. The study demonstrates that consumers whose behaviors were most exposed by Trended Data were more likely to open new accounts afterward.

Lenders thus stopped sharing the payment information the product relied upon. This eventually stemmed the flow of consumer switching, but it appears to have limited competition by making it more costly to acquire customers, the researchers write. It also helped to preserve the position of the incumbent lenders.

If actual credit-card payments information were still widely available and even used to generate credit scores, it would significantly improve the ability of these scores to accurately measure credit risk, the researchers calculate. The absence of this information hurts consumers by raising lending prices and limiting the amount of credit extended, not only on their credit cards but also on their auto loans and mortgages, according to the study.—*Neil Weinberg*

Benedict Guttman-Kenney and Andrés Shahidinejad, "Unraveling Information Sharing in Consumer Credit Markets," Working paper, March 2024. Go to **chicagobooth.edu/review** to read a longer version of this article.

The absence of payments information has consequences for consumers.

People in water-scarce areas think for the long run

Jonas Salk, who developed one of the first successful polio vaccines, once observed that "our greatest responsibility is to be good ancestors." His suggestion echoes the well-known Seventh Generation principle that some Native American tribes use to judge decisions by how they will affect our descendants far into the future.

One of the outstanding puzzles in social science is what disposes people to invest more thought in the future. Research by University of Queensland's Hamidreza Harati and Chicago Booth's Thomas Talhelm points to one possible factor: water access. People who live in water-rich areas tend to prioritize indulgence, while people who live in areas that suffer from water scarcity more often embrace a culture of conservation and long-term thinking, they find.

The work by Harati and Talhelm started with a comparison between two Iranian cities: Shiraz and Yazd. The cities are demographically and climatically similar, with one exception. Shiraz, known for its wine, is a garden-rich metropolis with plenty of rain. Yazd, 275 miles to the north, is one of the driest cities in Iran. Harati and Talhelm exploited this difference in two studies.

In the first, they recruited university students in the regions for a survey in which they asked respondents their thoughts about long-term orientation and indulgence. How much does persistence pay off, for instance? Or, how important is it to set time aside for having fun? The researchers find that students from water-scarce Yazd were inclined toward long-term thinking and downplayed the importance of indulgence; while students from Shiraz were the opposite.

For the second, a field study, the researchers posted job openings for a computer programmer, a customer-service agent, and an office assistant. There were two postings for each job, identical but for one attribute: whether they emphasized stability at a well-established company or excitement at a startup. The researchers find that students in

ILLUSTRATION BY JENNIFER TAPIAS DERCH

Yazd favored the stable job while those from Shiraz were more likely to apply to the startup.

A third study did not compare people from Shiraz and Yazd but instead primed students at the University of Tehran to think about water becoming more abundant or scarcer in the future. Students read scientific articles predicting that climate change would make water more abundant or scarcer. Using questions similar to those in the first study, they find that participants nudged to feel water is scarce reported more of a long-term orientation than those in the abundance group.

To generalize and reinforce the findings, Harati and Talhelm merged data on long-term orientation from the World Values Survey with per capita freshwater availability from the United Nations' Food and Agriculture Organization. Looking across nearly 100 countries, they find a consistent relationship between water availability and cultural disposition: more water meant more indulgent attitudes.

This work is part of a growing line of investigation in academic literature into the connection between environment and behavior—the way in which local ecology refracts and echoes in local culture. As the researchers note, there is something striking about the subtle but pervasive influence of people's environments. Though modern infrastructure that didn't exist previously allows the residents of Yazd to get water at their kitchen sink, the history of water scarcity persists in deep cultural currents.

A clearer understanding of the cultural origins of long-term orientation, Harati and Talhelm argue, is important given its relationship to many societal outcomes. One study of students in the same school system in Florida, conducted by University of Rochester's David Figlio, University of California at Los Angeles' Paola Giuliano, RAND Corporation's Umut Özek, and Northwestern's Paola Sapienza, finds that those whose parents were from cultures with a long-term orientation had better test scores, fewer absences, and higher graduation rates. Similarly, cultures that look ahead are known to save more money. Harati and Talhelm's data fit this narrative: people in Yazd saved 46 percent more than those in Shiraz in 2020 as a share of GDP. Yazd also has a higher literacy rate and a higher percentage of university-age residents attending college.

Finally, emerging evidence from Norwegian University of Science and Technology's Erik Saether, Ann Eide, and Øyvind Bjørgum suggests that groups that value the long run invest more in strategies to combat climate change. As Talhelm notes, there is hope in the fact that "the cultural value that helped humans adapt to environmental threats of our long-term past might help us adapt to the world's biggest environmental threat of the future."

–Dylan Walsh

Go to **chicagobooth.edu/review** to see citations for research mentioned in this article.

The economy looms larger than it used to in shoppers' decisions

Sanjay K. Dhar, James M. Kilts Jr. Professor of Marketing

Q1 How has retail marketing changed in the course of your career? If I go back to when I first started my research more than 30 years ago, we would often study what companies were doing in the grocery space—introducing a private label, merchandising a brand using different tools, innovating the retail format, using a different kind of coupon—to guide consumer behavior. But in the past decade, macroeconomic events have played a bigger role in the choices consumers make.

With that in mind, University of Washington's Shirsho Biswas, my Chicago Booth colleague Pradeep K. Chintagunta, and I recently looked at what happens to consumer purchasing behavior when household wealth, income, or employment status changes. This relates to earlier work we did on consumer responses to the Great Recession, and it helps to generalize those findings. Every economic shock is not the same—during a recession, your concern might be how you're going to put food on the table; during a pandemic, it could be more about staying healthy. So what patterns can we observe across macroeconomic events?

Brand managers allocate promotional dollars based on where demand is. On the retail side, category managers allocate funds based on what kind of brands people will be purchasing. Marketers need to be able to predict how economic shocks are going to affect what shoppers want.

PHOTOGRAPH BY JEFF SCIORTINO

Q2 **What is that research revealing?** Consumer responses are more nuanced than broad trends make them appear. For instance, we find that a 10 percent increase in income is associated with a substantial rise in spending on products that are not consumer packaged goods. Within CPG, that income shift results in dollars being taken away from the grocery store and going to warehouse clubs and discount stores, with more spending on national brands versus private labels.

It's generally accepted that when the economy is booming, national brands do well, and when it slows down, private labels do well. But our research demonstrates that in both cases, those brand types don't do well everywhere; they just do well on average. So the nuances of the interplay between product types, brand types, and store formats really matter for companies' decision-making.

Q3 **How should marketers be thinking about this shift in how the economy affects their market?** The way we want to think about it more broadly is that demand opportunities are created or taken away. For instance, COVID caused demand opportunities to be taken away from in-store purchasing and created for online shopping. Although the pandemic was a temporary event, that effect is not entirely reversible because people learned how convenient it is to buy things online, have groceries delivered, or order dinner through an app.

Demand opportunities are appearing and disappearing; new ways for value to be created and captured are emerging. All these changes are happening a lot more frequently today and have happened a lot more in the past 10 years than in previous decades. But it's cyclical. Right now we're going through a lot of turmoil—economic, political, and social—which means macro events are playing a bigger role. Once the dust settles, I think we'll go back to optimizing brand and retail strategies.

GENERATIVE AI CAN IMPROVE SCIENTIFIC EXPERIMENTS

ARTIFICIAL intelligence is rapidly changing jobs and industries, causing no small amount of consternation as it does. But on the bright side, it has the potential to greatly aid economists by, among other things, streamlining how they design and implement experiments, suggests research by the late Gary Charness, Chicago Booth principal researcher Brian Jabarian, and University of Chicago's John A. List.

Recent advances in generative AI, mainly through large language models, have sparked considerable interest. For one example, after OpenAI launched LLM-based ChatGPT, its valuation exploded, competitors rushed to keep up, and Microsoft kicked in $10 billion. Across the world, people are scrambling to understand how LLMs will transform jobs, the labor market, and various companies and sectors. (Read more in "AI is going to disrupt the labor market. It doesn't have to destroy it," in the Winter 2023/24 issue and online at **chicagobooth.edu/review**.)

Science, as many researchers have noted, is not immune. And as Charness, Jabarian, and List's paper explains, LLMs can help revolutionize how it is practiced. Addressing economists in particular, they write that LLMs could be harnessed to scale up experiments, make findings more accessible, and foster a culture of critical thinking about evidence-based analysis. LLMs could be used to improve nearly every step of an experiment, they explain—and they propose specific approaches for doing so. "All these offered directions require experimental benchmarking before becoming established scientific policies," qualifies Jabarian.

LLMs could be used to improve nearly every step of an experiment.

They group their recommendations into three categories: the design phase of an experiment, the implementation phase, and the analysis phase. LLMs could be used to analyze extensive data sets, identify gaps in knowledge, and help generate research ideas. AI could speed up the brainstorming phase while ensuring that research hypotheses are well-grounded.

Once a research question or hypothesis is in hand, LLMs could recommend a suitable experimental design, be it an economic game, market simulation, or something else. AI could help determine the optimal sample size for study.

In the implementation phase of an experiment, the real-time capabilities of LLMs become particularly useful, the researchers write. By functioning as interactive chatbots, LLMs could provide immediate support to participants, clarify instructions, answer questions, and ensure compliance with the experimental protocol.

And LLMs would significantly expand the scope and depth of data interpretation in the analysis phase, according to the research. Through techniques in natural language processing, they could analyze qualitative data such as participant feedback or chat logs and extract insights that traditional statistical methods might miss. They could organize and clean data efficiently, which not only speeds up the pre-analysis process but allows researchers to focus on interpreting results and drawing conclusions. And LLMs could be used to conduct statistical tests, generate visualizations, and identify patterns or correlations.—*Monika Brown*

Gary Charness, Brian Jabarian, and John A. List, "Generation Next: Experimentation with AI," *Nature Human Behaviour*, forthcoming. Go to **chicagobooth.edu/review** to read a longer version of this article.

How to get people to accurately predict price changes

People are bad at predicting inflation, which is a problem for policymakers charged with managing the economy because expectations about prices tend to be self-fulfilling.

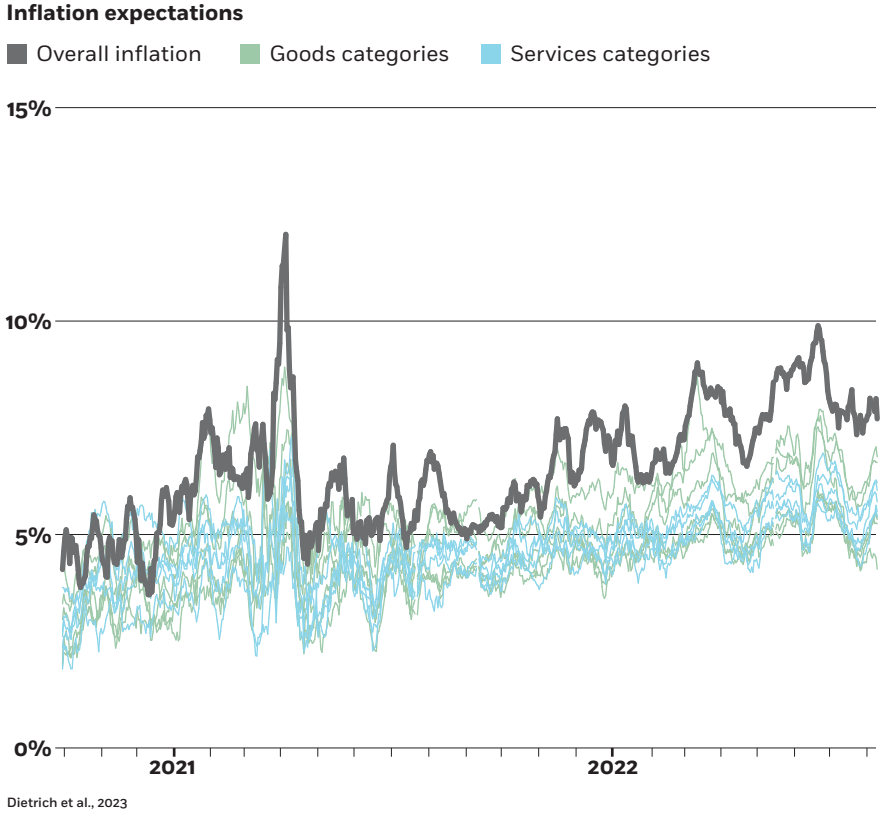
The Federal Reserve uses a monthly survey in which its New York branch simply asks nonexperts to forecast the Consumer Price Index. But there may be a way to get a more accurate fix on where consumers really think inflation is headed, according to a six-member international research team that includes Chicago Booth's Michael Weber.

Asking survey participants to forecast prices in 11 spending categories—rather than overall inflation—resulted in more realistic predictions, the researchers find. They conducted a two-year study involving almost 60,000 Americans. The results amount to a "proof of concept" for refining how central banks assess consumer sentiment, the researchers suggest.

Why do consumers' inflation expectations matter? If people believe that prices will rise sharply in the near future, they are more likely to buy things sooner rather than later, which may amplify inflation. It's a similar dynamic when workers demand higher wages because they expect prices to keep rising—but when wages rise, that drives up the cost of producing goods. Companies increase the price of those goods, and workers

A big versus small picture of inflation

US consumers' overall view of inflation was much higher than predictions for specific goods and services, according to a survey that asked Americans where they thought inflation was headed in the next year.



Dietrich et al., 2023

ILLUSTRATION BY PETE RYAN

then demand higher wages. Economists call this a wage-price spiral.

But consumers can be wrong and send false signals rippling through the economy. "Many consumers struggle to grasp the concept of inflation; they rely on salient cues when reporting forecasts; and survey responses are vulnerable to a host of cognitive biases," the researchers write.

While the concept of "inflation" in the aggregate can seem abstract, people may have a firmer sense of prices for specific types of goods and services, the researchers posited. To test their theory, in addition to asking about the overall CPI, they surveyed consumers about their expectations for prices of motor vehicles, recreational goods, other durable goods, food and beverages, gasoline, other nondurable goods, housing and utilities, health care, transportation, food services, and other services.

The researchers sent out questions via the Federal Reserve Bank of Cleveland's daily survey of consumers between July 2020, when inflation was extremely low, and August 2022, when it touched a 40-year high. They asked respondents what they thought the inflation or deflation rate in each category would be over the following 12 months. They also asked whether people expected to increase or decrease their own spending in those areas and by how much.

Respondents consistently said they expected overall inflation to be higher than the combination of their category-specific forecasts would signal. In fact, estimates for aggregate inflation were higher than estimates for any individual category. There were also sharper differences in estimates for overall inflation than existed in any category, the researchers find.

The survey responses yielded additional information about how consumers' backgrounds may play into their forecasts. For one, the primary grocery shopper in a household tended to have higher inflation expectations, and so did lesser-educated consumers and those from lower socioeconomic strata than more highly educated, higher-income respondents. Yet surveying consumers about individual components of inflation and combining the results always yields better aggregate forecasts than asking them about overall inflation, the researchers conclude.–*Andrea Riquier*

Alexander Dietrich, Edward S. Knotek II, Kristian O. Myrseth, Robert W. Rich, Raphael Schoenle, and Michael Weber, "Greater than the Sum of Its Parts: Aggregate vs. Aggregated Inflation Expectations," Working paper, November 2023.

FOR CONSUMERS, INFLATION HAS AN UPSIDE

INFLATION is typically thought of negatively, since rising prices erode purchasing power. But the trade-off is that it also erodes the real value of debt, and that can bolster household real wealth. Consumers who come to understand this dimension may change their outlook and their behavior, with broad economic implications, according to Goethe University Frankfurt postdoctoral scholar Philip Schnorpfeil, Chicago Booth's Michael Weber, and Goethe University's Andreas Hackethal.

They conducted a randomized controlled trial involving more than 3,000 customers of a German bank in July 2022, when inflation was at a 70-year high of 8.7 percent. The researchers first asked participants a series of questions about the economy, their assets, and estimated changes in their net wealth over the previous 12 months. They then split them into a control group, which received information only about the rate of inflation; a "savings-erosion" group that read about how inflation could hurt savers; and a "debt-erosion" group that learned how inflation benefits debtors.

Then they asked participants about their expectations over the following 12 months for real-estate prices, unemployment, interest rates, and household income. They also surveyed the participants about their planned spending.

The researchers find an "asymmetric awareness of the erosion channel" of inflation. Three-quarters of all participants knew that inflation would reduce the real value of savings, but only a third understood it would also lower the real value of fixed-interest-rate debt. The group that learned about inflation's impact on savings subsequently was more bearish about savings than the control group. Those who received information about debt erosion were less credit averse.

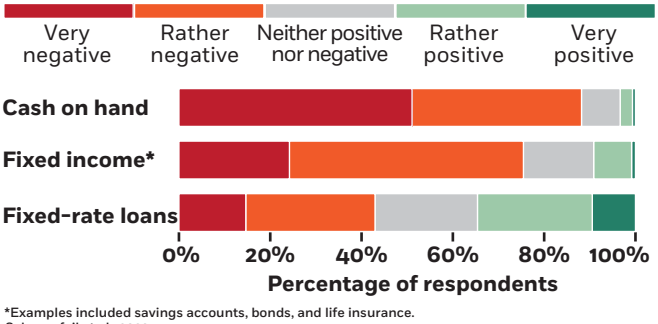
The researchers followed the participants for several weeks after the experiment. Those in the debt-erosion group increased their planned and actual spending.—*Andrea Riquier*

Philip Schnorpfeil, Michael Weber, and Andreas Hackethal, "Households' Response to the Wealth Effects of Inflation," Working paper, September 2023. Go to **chicagobooth.edu/review** to read a longer version of this article.

But what about debt?

While most study participants were aware that inflation hurts the real value of cash and fixed-income investments, only a third understood that it helps borrowers of fixed-rate loans.

Perceptions about the wealth effects of inflation, by type of financial instrument



*Examples included savings accounts, bonds, and life insurance.
Schnorpfeil et al., 2023

WHAT'S A TRADEMARK WORTH?

For the first time, researchers have come up with a figure—$22.5 million for the median trademark, according to Nova School of Business and Economics' Pranav Desai, Ekaterina Gavrilova, Rui Silva, and Margarida Soares. They modeled a new system for valuing trademarks based on the stock market reaction to individual trademark registrations. Their analysis tracked outcomes for publicly traded companies from 1961 through 2021 for trademarks registered up until 2016.

The researchers also used NielsenIQ Retail Scanner Data from Chicago Booth's Kilts Center for Marketing to study how trademarks affect product dynamics. They compiled data on sales, quantities, and prices for each product in the database between 2006 and 2020.

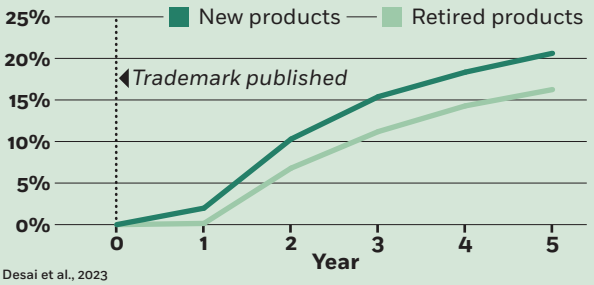
They find that companies that publish trademarks tend to subsequently launch more new products, expand sales, invest more in physical capital, hire more people, increase production, generate more profits, and take greater market share. Trademarks were also associated with more product obsolescence, which suggests companies filed trademarks as they refreshed their offerings.—*Francine McKenna*

Pranav Desai, Ekaterina Gavrilova, Rui Silva, and Margarida Soares, "The Value of Trademarks," Working paper, October 2023.

Out with the old, in with the new

Trademark publication was associated with a nearly 21 percent rise in product launches and about a 16 percent increase in product retirements.

Product category outcomes after trademark publication
In response to a 1 standard deviation increase in trademark output



Desai et al., 2023

Why any change to the goodwill rule is a big deal

Accountants have been wringing their hands for decades over how to treat a business's goodwill—the value of customer loyalty, human capital, and synergies—when the company changes hands. Should it be allowed to sit on the books perhaps indefinitely, or should an acquirer write it off over a defined period of time?

The issue has been a hot potato for years, and research explains why the stakes are high: Currently the carrying value of goodwill is tested each year for impairment to determine whether its fair value has decreased. Changing the rules to require that buyers amortize, or gradually expense, goodwill over 10 years would slash buyout prices, dramatically shrink the $1.6 trillion-a-year US mergers and acquisitions market, and push more businesses into the arms of private-equity buyers, according to Rice University's Stefan J. Huber and Chicago Booth's Charles McClure.

Goodwill, formally defined as the difference between the purchase price of a business and the fair value of the net identifiable assets acquired, typically accounts for almost half of US corporate deal values, according to the researchers. As the largest share of the purchase price paid in corporate mergers, it can significantly affect the amount an acquirer is willing to pay, they write.

The Financial Accounting Standards Board, the 51-year-old private standard-setting organization for US companies, had not long ago been considering a change for the accounting of goodwill. For almost 20 years, it has held that goodwill isn't subject to amortization but should be tested annually for impairment that could force a write-down. Research in 2012 by Oxford's Karthik Ramanna and MIT's Ross L. Watts suggested that the FASB's 2001 ruling "was the result of political pressure" from businesses that preferred the impairment approach, Huber and McClure write.

In 2018, FASB revisited the issue and considered whether to revert back to requiring companies to amortize goodwill over 10-25 years. It dropped the idea in 2022, with FASB chairman Richard Jones citing the magnitude of the change and uncertainty about the impact.

To better understand that impact, Huber and McClure built a model based on 861 all-cash deals involving takeover auctions of publicly traded companies from July 2001 through September 2022. They make a distinction between strategic bidders—such as competitors, customers, or suppliers seeking to increase earnings—and financial bidders, such as private-equity funds, which prioritize maximum cash flow.

In each case, the bidders' valuations of the target depend on the resulting mix of earnings and cash flow from the merger. Strategic bidders are willing to pay more for companies that produce earnings under the current approach, which involves reducing the value of goodwill assets only when conditions change rather than consistently over a fixed period.

The researchers find that strategic bidders paid more than financial bidders under the current goodwill accounting rules. They calculate that if the rules were changed to require buyers to expense goodwill over time, all bidders' premiums over the market price would fall by 6 percentage points, and M&A volume would decline by 4.29 percent, or $68.6 billion a year. Premiums offered by strategic buyers would fall by 13 percentage points, and the proportion of private-equity acquirers would rise by 7.74 percentage points.

"Such changes in the makeup of winners can influence the ownership of a substantial portion of the economy," the researchers write. "Adopting an accounting standard that amortizes goodwill reduces the relative strength of strategic bidders" but "does not affect financial bidders' values." They estimate that such a shift would increase the likelihood of a financial bidder winning the takeover from nearly 30 percent to about 37 percent.

The findings could be of interest to the US Securities and Exchange Commission, which has expressed concern about the growth of PE ownership at the expense of public investment. Accounting standards, note the researchers, "can contribute to the balance between public and private markets." Also, intangible assets such as goodwill have become a significant source of economic value. Goodwill assets alone on the balance sheets of S&P 500 companies increased 37 percent from 2017 to 2022, when M&A activity was booming, according to data provider Calcbench.

McClure and Huber say that their paper, by quantifying the economic consequences of different accounting treatments, can serve as a guidepost for standard setters, who "continue to debate how to account for intangible assets."—*Francine McKenna*

Go to **chicagobooth.edu/review** to see citations for research mentioned in this article.

Learning that a company's history included black leadership made a difference.

CELEBRATING CORPORATE HISTORY CAN BACKFIRE

MANY COMPANIES and other organizations, from religious institutions to universities, celebrate their history. It can be a means of engaging with potential employees, as well as customers and others. But Black Americans may experience an organization's celebration of its history as marginalizing and even threatening, find Chicago Booth postdoctoral scholar Laura E. Wallace, WGU Labs' Stephanie L. Reeves, and Ohio State's Steven J. Spencer, who write that the reaction is related to the fraught racial history of the United States.

Focusing on the workplace, the researchers explored the idea of a social-identity threat, the fear that a person might be devalued or excluded on the basis of her group. "Members of marginalized groups are often hypervigilant for environmental cues that they will be treated differently based on their group membership," Wallace, Reeves, and Spencer write.

Could old photographs be interpreted as such a cue? The researchers performed an online experiment in which Black American participants evaluated the website of a fictional consulting company, rating whether they felt it would value them as employees and indicating how likely they would be to apply for a job there. All of the participants viewed the company's About page, which showed a black-and-white photograph of four white male founders, alongside a caption stating that it had been taken in 1951 in Charleston, South Carolina.

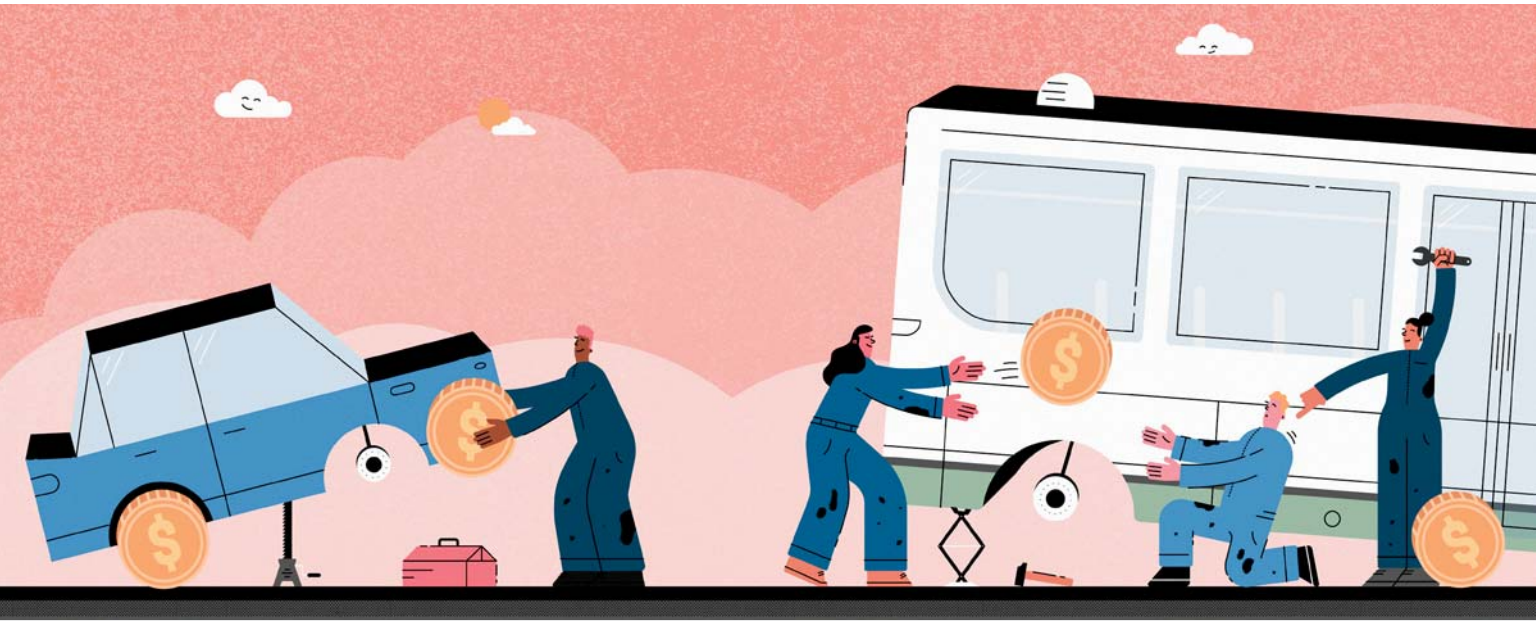
"We expected that participants would interpret the photo of the all-White founders and the information that the company was founded in the South during the Civil Rights Era as an indication that Black Americans were historically marginalized in the company," the researchers write.

Half the participants were shown a version of the page in which the company celebrated its history, while the other half viewed a version where history wasn't discussed beyond the photograph. As the researchers predicted, participants who read the company's history said they anticipated feeling less of a sense of belonging in the organization and reported less of an intention to apply for a job there, compared with those who didn't read the history.

In another experiment, the researchers wanted to see whether a description of a company's history would have the same effect if accompanied by fewer additional cues. Even without the photo, Black participants who read about the company's history anticipated less of a sense of belonging in the organization and expressed less intention to pursue employment. And, in another experiment, it was true even when mention of the company's Southern founding location was removed.

However, learning that a company's history included Black leadership made a difference. In a final experiment, participants who were told that Black leaders had been part of the company's past, compared with those who weren't told so, felt more of a sense of belonging and reported greater intention to apply for employment.—*Alice G. Walton*

Laura E. Wallace, Stephanie L. Reeves, and Steven J. Spencer, "Celebrating Organizational History Triggers Social Identity Threat among Black Americans," *Proceedings of the National Academy of Sciences*, February 2024. Go to **chicagobooth.edu/review** to read a longer version of this article.



Charge drivers to improve public transit

Cars create plenty of problems, from road congestion to climate-damaging carbon emissions to unhealthy tailpipe pollutants. But cars are convenient, part of the reason public transit accounts for only 3.4 percent of the 850 million trips taken in US cities every day.

But as demonstrated by New York's failed congestion pricing initiative, shifting the balance away from cars in an economically fair and efficient way is not a straightforward task for cities. That said, an analysis of Chicago's transit system by Chicago Booth's Milena Almagro, University of Pennsylvania PhD student Felipe Barbieri, Penn's Juan Camilo Castillo, MIT PhD student Nathaniel Hickok, and MIT's Tobias Salz finds that blending charges for drivers with adjustments to transit fares and schedules offers a promising path.

"If we think about urban transportation, it's important to think about a combination of policies at once," Almagro says. "One policy can serve as a complement to others. In this case, road pricing collects money that is reinvested in public transit."

The researchers considered three transportation policies: adjusting bus and train fares, changing bus and train frequencies, and introducing a fee on cars driving into the city.

They compiled data for January 2020 from a range of sources. These data included the station of origin, time of day, and inferred drop-off location for every trip taken on buses and trains.

The researchers also analyzed every taxi and ride-share trip, looking specifically at pickup and drop-off points, price, number of riders, and trip length and duration. And they used anonymized cellphone data to track commuters from home to work and back. Finally, they matched the information with demographic data from census tracts, inferring the income levels of commuters.

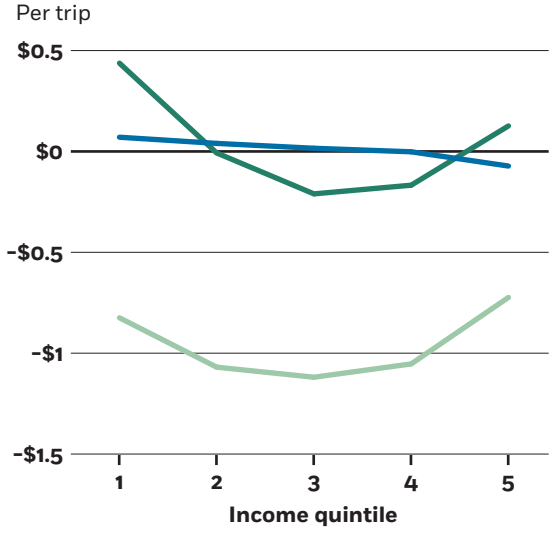
The researchers modeled several scenarios, starting with a world in which the Chicago Transit Authority has an unlimited budget. In this case, rides would be free and trains would run more frequently. (Buses would actually run less often to optimize bus ridership, as most buses operate well below capacity.) On the basis of the current CTA budget, this scenario would create a deficit of $11 million a week.

How the policies would affect different income groups

Low-income consumers would gain the most from lawmakers combining a public-transit policy with a road tax and a rebate.

Change in consumer surplus across income quintiles, compared with the status quo

- **Public transit with a budget constraint:** The city reduces bus and train fares, but also reduces frequencies.
- **Road tax:** The city charges a road tax. Transit prices and frequencies don't change.
- **Public transit + road tax with a rebate:** The city uses a portion of the revenue to reduce transit prices and increase frequencies, and rebates the rest to consumers.



Almagro et al., 2024

ILLUSTRATION BY ERRATA CARMONA

More realistic scenarios would be making changes to public transit or adjusting road pricing, but these options taken alone would create distinct problems, the researchers find. Because the CTA has tightly constrained resources, it could by itself make only modest changes to fares and train frequencies. It also has considerations to manage beyond cost, such as offering services equitably. These factors contribute to a large gap between how well the CTA could be serving its customers and how well it actually serves them.

The dilemma with using road pricing on its own is that the cost is unevenly distributed. While these tolls could effectively reduce the number of cars on the road and generate environmental benefits, the researchers find that under their optimal price—an average of $13.40 daily per driver—commuters would pay nearly $30 million a week. People in the middle class, who make up the largest share of car commuters, would bear most of the expense. Lower-income groups tend to take public transit, especially buses; wealthier people tend to walk, travel from areas with good access to trains, or use a ride-hailing service.

If authorities pursue both policies simultaneously, the researchers find, the revenue collected from charging commuters 30 cents a kilometer could subsidize cheaper, more convenient public transit. The CTA could slash fares to next to nothing: 16 cents for a bus fare compared with $2.25 today and 26 cents for the train, down from $2.50. Excess revenue from road fees (any collected beyond the amount required to finance the CTA at this level) could go back to residents in the form of rebates, lessening the budgetary strain on middle-income commuters.

Almagro notes that the specifics of transit infrastructure and its use vary widely from city to city. The system of trains and buses in Chicago differs from that of New York or Los Angeles. The number of commuters and their transportation patterns also differ. But the central finding is universal, she says.

"Road pricing seems to be the most effective instrument when thinking about efficiently reducing congestion and environmental externalities, but it's also the one that hurts consumers the most," Almagro says. "The question then becomes how road pricing can best be combined with investment in public transit to make everybody better off."—*Dylan Walsh*

Milena Almagro, Felipe Barbieri, Juan Camilo Castillo, Nathaniel Hickok, and Tobias Salz, "Optimal Urban Transportation Policy: Evidence from Chicago," Working paper, May 2024.

More competitive arbitrage could increase the risk of runs, research finds.

WILL STABLECOINS DESTABILIZE OTHER MARKETS?

STABLECOINS may not be as stable as intended, suggests research by Columbia's Yiming Ma, University of Pennsylvania's Yao Zeng, and Chicago Booth's Anthony Lee Zhang. They identified the risk of runs in the two biggest dollar-backed stablecoins as "economically significant" in an analysis of data covering two recent years.

Stablecoins, a form of cryptocurrency, are pegged to the US dollar and backed by assets such as US Treasury securities and corporate bonds. But like money market funds, stablecoins can fall below $1 and force their sponsors to sell assets in a potentially market-destabilizing fire sale.

In the United States, Federal Reserve chair Jerome Powell urged Congress last year to impose robust federal regulation of stablecoins. (It hasn't done so.) The value of stablecoins worldwide surged to more than $130 billion by the beginning of 2022 from $5.6 billion two years earlier, the researchers report.

Ma, Zeng, and Zhang analyzed the role of arbitrageurs, who buy and sell stablecoins in response to fluctuating demand to keep their value constant.

They collected transaction-level data on each stablecoin creation and redemption for the six largest dollar-backed stablecoins. The amount of data collected varied by coin. The researchers also obtained trading prices from the main crypto exchanges and data on the reserve assets of the two top stablecoins in terms of transaction volume, Tether and Circle, at various points in 2021 and 2022.

Their analysis reveals that only a handful of arbitrageurs were able to redeem stablecoins for $1 in primary markets. The small number of such market players surprised the researchers because more participants would improve market efficiency.

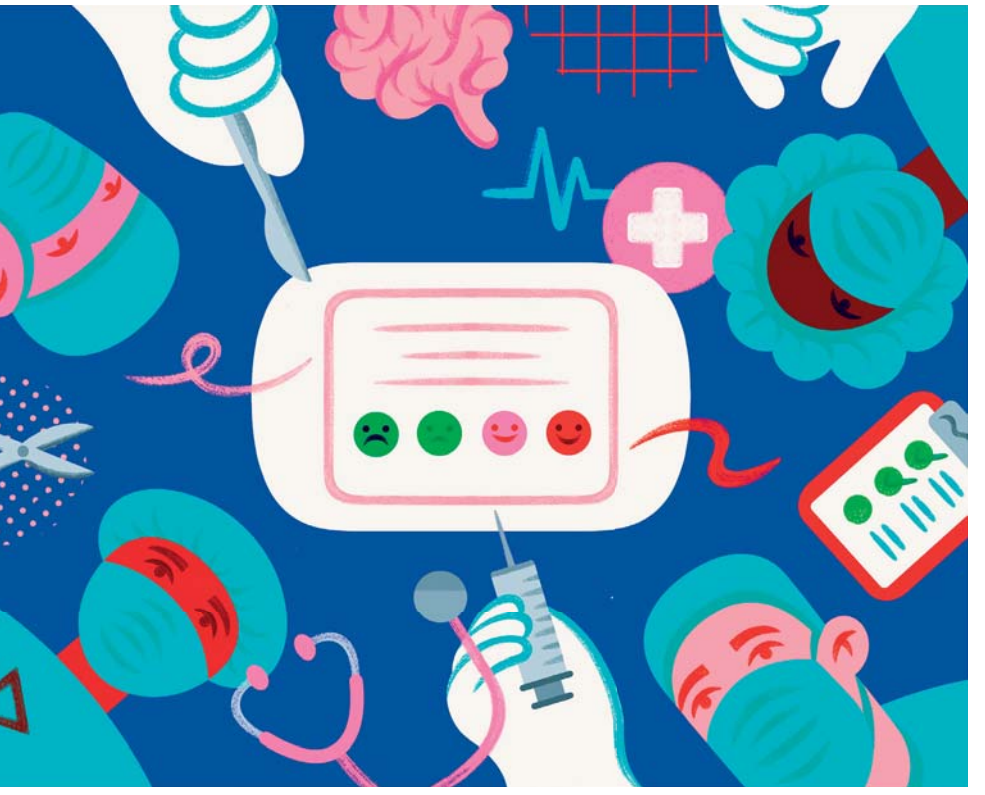
However, they find that issuers face a tradeoff: while efficiency is generally desirable, more competitive arbitrage could increase the risk of runs. "Stablecoins are subject to panic runs because of illiquidity in their assets and the fixed $1 redemption value," they write. Yet more efficient arbitrage would just make it easier for investors to sell, according to their results. The researchers also demonstrate that market prices frequently deviate from $1, and that doesn't necessarily trigger a run.

The researchers plugged data on the Tether and Circle reserve assets into their model to put a figure on the probability of a run on either of those. They find that Tether's assets were less liquid than those of Circle. As of September 2021, the risk of a run amounted to 2.5 percent for Tether and 2.1 percent for Circle, the researchers calculate.

A run on a major stablecoin could have negative implications for debt markets, the study suggests. If Tether had to sell its Treasury positions in a run, the researchers write, that would amount to one-sixth of the amount of Treasury securities that mutual funds liquidated in the March 2020 COVID-inspired global dash for cash.

Ma, Zeng, and Zhang note several policy implications of their findings, including that regulators should pay close attention to stablecoin arbitrage capacity, and coin issuers and regulators could reduce the risk of runs by imposing redemption fees on arbitrageurs. In addition, their model predicts that Tether and Circle could meaningfully reduce their run risks and increase stability by paying dividends to investors. —*Michael Maiello*

Yiming Ma, Yao Zeng, and Anthony Lee Zhang, "Stablecoin Runs and the Centralization of Arbitrage," Working paper, March 2024.



A way to improve teamwork in operating rooms

In an operating room, where human life can hang in the balance, communication among members of a surgical team is crucial. But it's hard to assess how well those team members work together.

The lack of data about this is due in large part to the high-pressure environment, where it's infeasible to implement conventional survey tools such as questionnaires, write Yale's Kiran K. Turaga, University of Chicago's Hunter D. D. Witmer, and a team of researchers. They demonstrate a new way to collect real-time, qualitative data from OR staff—and their method could potentially improve surgical outcomes.

The project was conducted in the Healthcare Analytics Laboratory at Chicago Booth, led by the team's Dan Adelman. Over a 24-week period in 2021, the team collected feedback at 30 University of Chicago Medicine operating rooms. The researchers installed 120 Smiley Terminals manufactured by the HappyOrNot company. They equipped the ORs with four terminals, one for each member of the surgical team: surgeon, anesthesiologist, scrub nurse, and circulator (the nurse responsible for ensuring that surgeons have everything they need to perform a procedure). After completing a procedure, team members were asked to assess the teamwork on a 4-point scale: very satisfied, satisfied, dissatisfied, and very dissatisfied. On the terminal, there was a happy or a frowny face associated with each rating.

The researchers collected about 4,100 responses from about 2,100 OR teams. Anesthesiologists were the most satisfied—93 percent of them clicked the smiley faces associated with either "satisfied" or "very satisfied"—followed by surgeons, at 88 percent, and the two types of nurses, at 80 percent or more, the researchers find.

The results indicate a high level of satisfaction with teamwork quality, but the study also has limitations. The overall response rate was only about 32 percent and varied considerably according to the type of surgery. Nearly two-thirds of team members participating in cardiac operations responded, but fewer than a tenth of those in trauma surgeries did. Because no high-performance benchmarks exist, it's impossible to know how teamwork on these surgical teams stacked up against surgical teams elsewhere.

The paper's primary contribution is proof of concept for continuous qualitative data collection in an OR setting, Turaga says. He pointed to two additional findings as especially important: A higher degree of specialization among scrub nurses was associated with greater overall satisfaction. And when satisfaction with teamwork was highest, quality of care was better.

To get at the role of specialization, the researchers broke out team member feedback by service line—cardiac surgery, neurosurgery, or trauma surgery, for example. Scrub nurses and circulators were assigned an experience score based on the number of surgeries in a given service line they'd assisted with in the previous six months. When scrub nurses and circulators had a higher degree of specialization, other team members were more likely to indicate they were "very satisfied" with the quality of teamwork.

In cases where the level of satisfaction with teamwork quality achieved this rating, there was a 15 percent reduction in the length of a patient's hospital stay. Hospitals commonly use length of stay as a proxy for quality of care because it indicates there were fewer complications or inefficiencies.

Specialization can improve teamwork by making the generally consistent steps of a surgery even more seamless, Turaga says. "There can be almost no talking between a surgeon and a nurse, or a surgeon and a circulator," he says. "There are times when a surgeon can just reach out a hand, and the other person knows exactly which instrument the surgeon needs without exchanging a word. In these situations, they know what equipment is needed, and the exact amount of equipment needed."

At the same time, Turaga notes, healthcare institutions seldom take the steps necessary for building OR teams with the level of specialization needed to perform at the highest level. While surgeons and anesthesiologists must be specialized, nurses often do not receive specialized training for assisting in many types of surgery. Turaga argues that the research findings make the case for a higher degree of specialization in OR teams. This would improve satisfaction, teamwork, and, ultimately, patient care.—*Ty Burke*

Hunter D. D. Witmer, Joshua A. Morris-Levenson, Çağla Keçeli, Frederick A. Godley IV, Ankit Dhiman, Dan Adelman, and Kiran K. Turaga, "Novel Application of a Dynamic, In-Room Survey Platform to Measure Surgical Team Satisfaction," *Annals of Surgery*, January 2024.

ILLUSTRATION BY CHANELLE NIBBELINK

US bank regulators could have averted $9 billion in losses

The failures of Silicon Valley Bank and First Republic Bank in 2023 caused plenty of finger-pointing at bank regulators. Many US financial institutions including SVB and First Republic were covering large portions of their liabilities with assets such as government bonds that stood to lose value if interest rates rose significantly–which is exactly what happened. As the Federal Reserve hiked interest rates, rumors swirled that the two banks might not have sufficient assets to cover all their deposits. Customers yanked their money out, causing runs at both.

Indiana University's Yadav Gopalan and Chicago Booth's João Granja find through some rough calculations that had regulators acted two quarters earlier, they would have averted $9 billion in losses. However, the researchers also say their analysis of the events reveals a more complicated picture than examiners being oblivious to systemic risks–plus it's particularly hard to assess regulators' performance given how necessarily secretive their work is. "Regulators may receive the blame for some very public failures but no glory for the bank failures that they prevented from happening," write Gopalan and Granja.

To understand what transpired, the two researchers matched information maintained by banking regulators with publicly available data about banks' financial conditions. The regulatory data, which cover late 2020 through early 2023, show the risk scores supervisors assigned to the banks, while the publicly available data include factors such as interest-rate risk exposure, exposure to uninsured deposits, bank size, bank capitalization, and asset quality.

The onset of interest-rate hikes in early 2022 served as a natural laboratory for assessing how bank examiners managed a credit shock. "In our setting, a well-defined event, the Federal Reserve's decision to raise interest rates, triggered a shock to the value of equity of banks with significant maturity mismatches and unstable deposits," write the researchers.

For each financial institution, regulators issue what's known as a CAMELS rating. The acronym refers to the six categories that are assessed: capital adequacy, asset quality, management, earnings, liquidity, and sensitivity to risk. The researchers find that as rates rose, regulators adjusted the ratings they issued to the most-exposed banks, downgrading two of the six risk factors the most: liquidity and sensitivity to risk. A rate shock naturally draws more attention to these components than the others, so the downgrades suggest that supervisors "understood the consequences" of interest-rate risk, the researchers write.

Examiners were less likely to downgrade banks that had derivative contracts that protected them against rate movements, another indication that they understood the emerging risks.

Regulators might have been slow to scrutinize balance sheets because of an accounting tactic. According to other research by Granja, troubled banks were more likely to classify securities as being "held to maturity" rather than "available for sale," perhaps to avoid marking down the value of those assets. (For more about this research, read "Are US banks hiding their losses?" in the Spring 2024 issue and online at **chicagobooth.edu/review**.) The stalling tactic might have worked, for a while. Starting in the second quarter of 2022, examiners were more likely to downgrade banks with larger unrealized losses in AFS securities, but only began to significantly downgrade banks with HTM portfolio losses during the first quarter of 2023.

Finally, regulators were not more likely to downgrade banks that relied more heavily on uninsured deposits, suggesting that they did not fully appreciate the risks associated with this type of deposit. Examiners, investors, and bank executives have long assumed that deposits are "sticky," meaning that customers won't pull their money out until they absolutely have to. But online banking makes it much easier to move funds, even just to pursue higher interest rates on deposits. (Read more in "Why your banking app might spell trouble for your bank," Spring 2024 and online.)

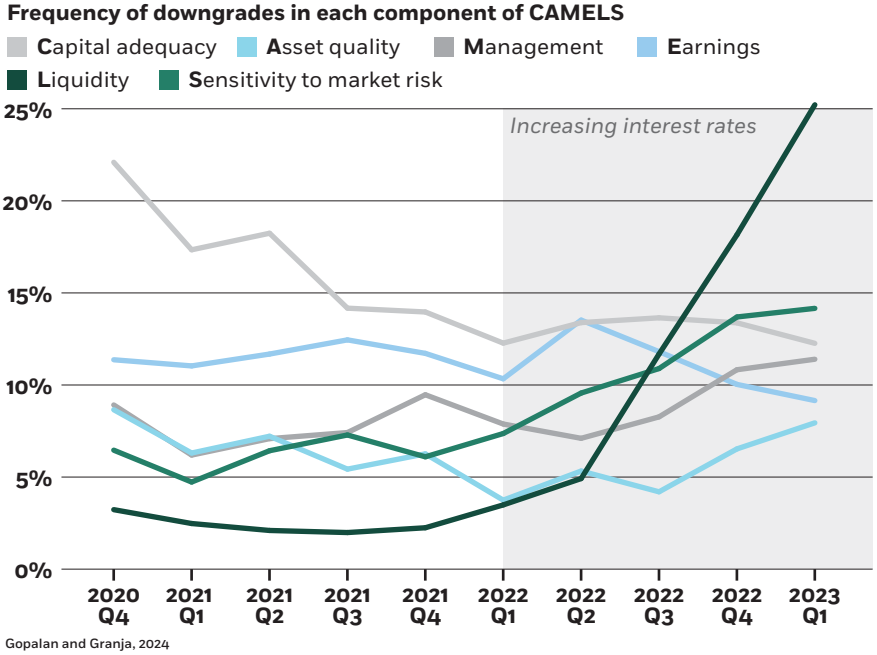
"This inaction of supervisors may be surprising," Gopalan and Granja write. But, they say, it also may confirm supervisors' own admissions in the wake of the fallout of SVB that current supervisory models don't go far enough to capture the additional liquidity risks associated with an unstable deposit base, including uninsured deposits.

It's clear that regulators did catch and contain some of the risks, saving money for taxpayers and bank customers alike, the researchers conclude. But considering the steep costs of delayed action, it's fair to question whether the failure to detect a fairly straightforward risk–and move quickly to stop it–indicts the entire supervisory system and the $2 billion a year it costs taxpayers, Gopalan and Granja write.–*Andrea Riquier*

Go to **chicagobooth.edu/review** to see citations for research mentioned in this article.

Some saw the risks

As interest rates rose, US financial regulators assessed financial institutions using a CAMELS rating. They downgraded its "liquidity" component more than the others.



Gopalan and Granja, 2024

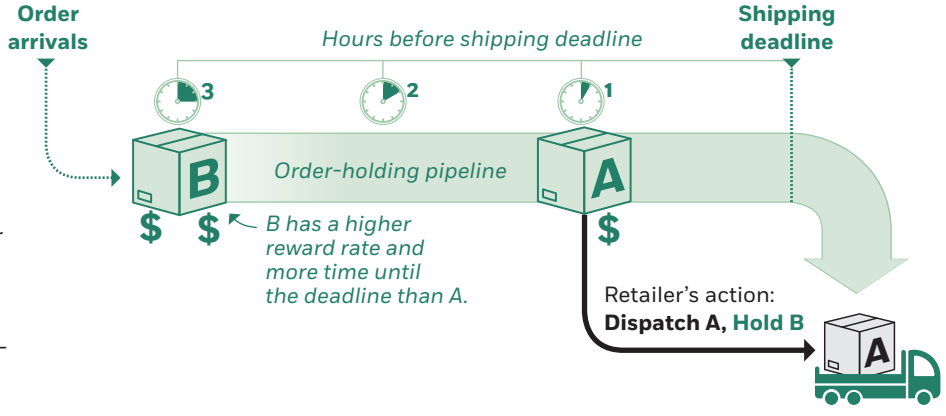
At the e-commerce warehouse, a distribution dilemma

Online shoppers often place two or more orders in quick succession—an initial one, and then a follow-up one when they remember something else they meant to buy. To counter the rising costs this multiordering creates, some e-commerce platforms hold orders before fulfilling them, but this welcomes its own complications. Chicago Booth PhD student Mohammad Reza Aminian, Columbia's Will Ma, and Booth's Linwei Xin built a decision-making model to determine in real time which orders companies should hold.

Some orders are straightforward to fulfill. . . .

Imagine a retailer that can hold one order at a time for a maximum of three hours. As another customer's order comes in, the retailer must decide whether to continue holding the first order or dispatch it in favor of delaying the new one.

Scenario 1: An order for **Customer B** arrives. At the same time, the retailer is holding an order for **Customer A** that has one hour left on the clock. On the basis of each customer's probability of reordering and their spending patterns, the retailer predicts that its **"reward rate"** from holding **Customer B's** order is higher than that of holding **Customer A's.** The reward rate represents the potential savings per unit of time on shipping costs, which are proportional to a customer's multiorder rate.



Scenario 2: An order for **Customer D** arrives while the retailer is holding **Customer C's** order. Although **Customer C's** order has only 1 hour to go before the dispatch deadline, while **Customer D's** order would have three hours to go if the retailer decided to delay this order instead, the reward rate from holding **Customer C's** order is much higher than that of holding **Customer D's.**

. . . But other orders present a trade-off.

Scenario 3: The retailer is holding **Customer E's** order when **Customer F's** order arrives. The reward rate from holding **Customer E's** order is slightly higher than that of holding **Customer F's**, but **Customer F** has more time left before the dispatch deadline, which could tip the scales.

Mohammad Reza Aminian, Will Ma, and Linwei Xin, "Real-Time Personalized Order Holding," Working paper, November 2023.

Three algorithms for holding orders

Orders that have a high chance of consolidation and those that have more time in the system before being dispatched are good candidates for holding. But these criteria can conflict, creating a trade-off. When this happens, retailers can follow one of three algorithms:

Reward-rate algorithm
Hold the orders with the highest reward rate.

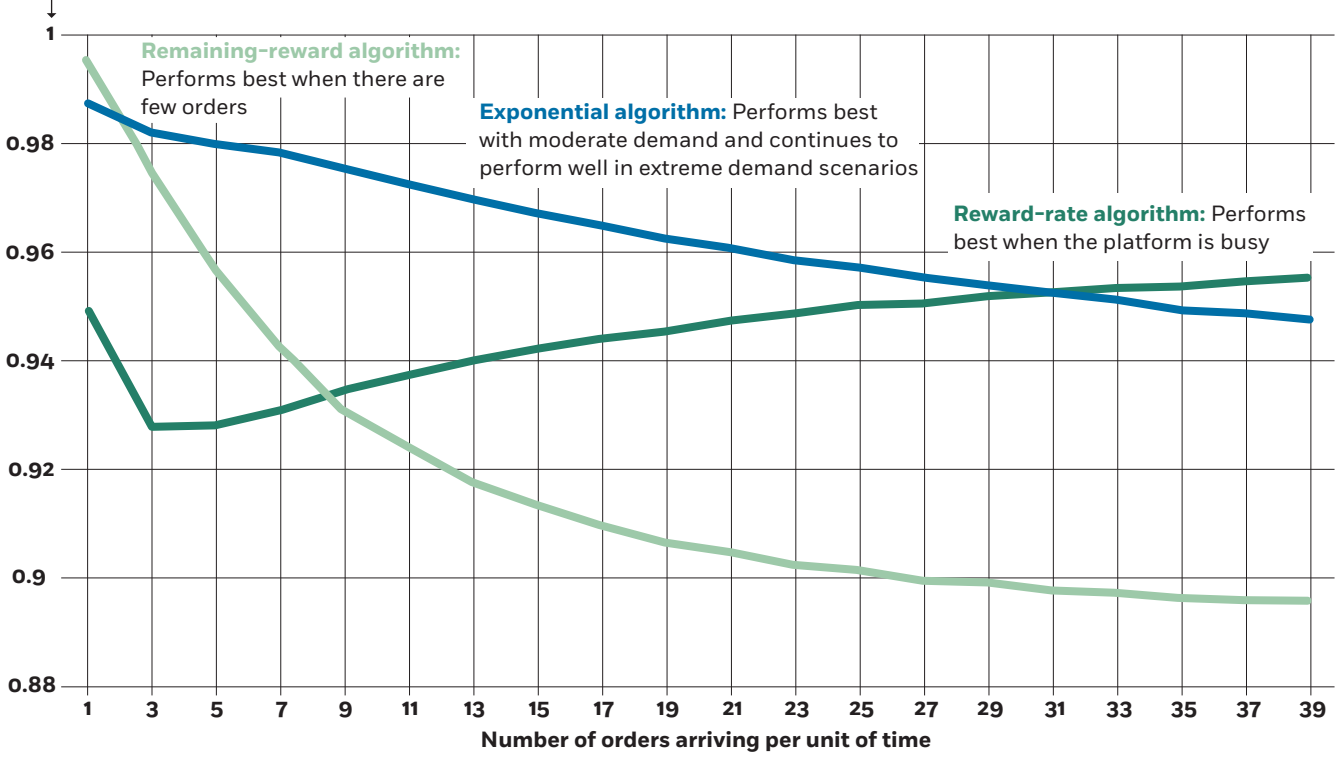
Remaining-reward algorithm
Hold the orders with the largest remaining reward, which is the reward rate multiplied by the remaining time to dispatch.

Exponential algorithm
Hold the orders with the largest value from an equation that considers both the reward rate and the remaining reward (calculated by multiplying the reward rate by an exponential function of the remaining time to dispatch).

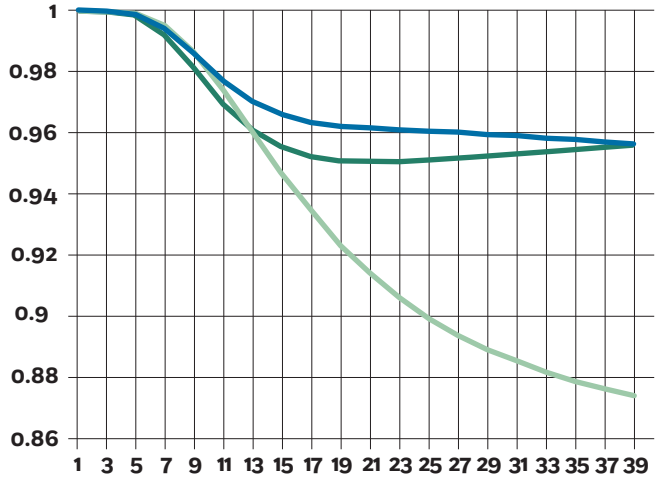
The best option depends on the retailer's capacity constraints and the platform's busyness. But in most cases, the exponential algorithm generates the best performance relative to a benchmark. Here's how the algorithms perform among platforms with different holding capacity:

Platform can hold only one order at a time

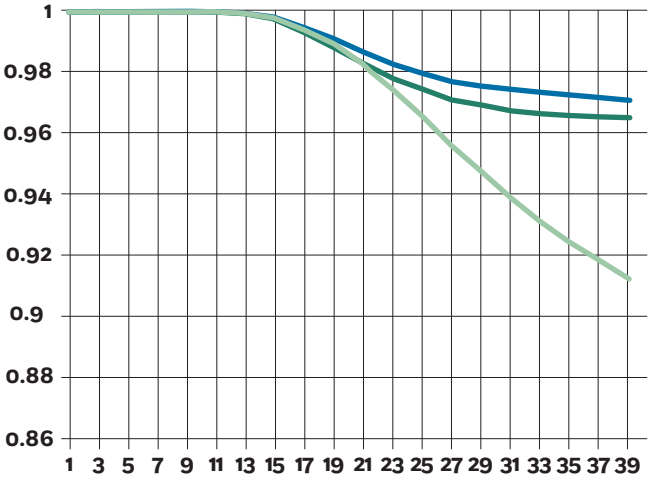
Ratio of an algorithm's average performance to that of a benchmark algorithm that knows all future customers in advance



Platform can hold 10 orders at a time



Platform can hold 20 orders at a time



(THAT YOU MIGHT NOT)

Ten ways investors are, or should be, using large language models

STORY BY BY MONIKA BROWN
ILLUSTRATION BY MADE UP STUDIO

Great investors tend to be avid readers, always hunting for some piece of information to give them a financial edge. There are decades if not centuries of examples of pros who have combined something they've read—in a book, article, or regulatory filing—with their market experience to gain a lucrative insight. For one example, investment manager Jim Chanos's careful reading of Enron's regulatory filings, and his past experience with fraud detection, led him to suspect accounting irregularities at the company. He made $500 million when Enron filed for bankruptcy in 2001.

These days, though, even the most avid readers would have trouble competing with the volume of financial insights that artificial intelligence, in the form of large language models, can uncover. LLMs have gained mainstream popularity thanks to OpenAI's ChatGPT, an advanced chatbot powered by a series of generative pretrained transformer language models. OpenAI has released several versions of its LLM, with GPT-3.5, GPT-4, and GPT-4o among the most recent.

Almost a decade ago, *Chicago Booth Review* published a feature titled "Why words are the new numbers" about a coming revolution in text analysis. That predicted revolution arrived, and it demolished the monopoly that numbers long held in forecasting models. Numbers are still important, of course—but text analysis is ascendant and everything is now potential data.

The candid speech during earnings calls? Data. The formal prose of annual filings? Data. News articles? Data. The entire internet? Data.

LLMs are trained on vast amounts of text covering a broad range of information and can apply their repositories of knowledge to evaluate new information. Where a human will depend on past experience and intuition, LLMs use data and patterns from their training.

And they operate at a scale that exceeds human capabilities, quickly analyzing mountains of text and allowing traders and investors to mine insights faster and more accurately than was ever possible. They can connect ideas from different parts of a text to create a better understanding of its overall content. LLMs can even be customized, trained to become experts on accounting irregularities—or, say, mall leases or risk management.

Every asset manager with a technology team now has the opportunity to wield—and profit from—an enormous knowledge base, and many are doing just this. Funds are using LLMs to read and glean insights from earnings call transcripts, 10-K regulatory filings, annual reports, social media, and streaming news headlines—searching for clues about a company's direction.

From the output of this text mining, LLMs can create direct trading signals (instructions to buy or sell) or develop new predictive variables for their forecasting models. If you hold actively managed funds in your retirement accounts, there's a good chance the pros running the strategies are harnessing the research power of LLMs.

It makes sense to ask whether the advantages of LLM strategies will disappear as soon as everyone else uses them too. That's been the outcome with arbitrage strategies—their returns fall when too many investors are chasing the limited opportunities. However, the opportunities here appear more bountiful than in arbitrage scenarios. With the field in its early stages, researchers are still finding new ways to apply AI to tease out investment insights and trading opportunities. Plus, new data sources that run the gamut from text to image, audio, and video are enabling the uncovering of information that is not so easily priced into the markets. (See "Images and audio are now data too," page 38.)

Researchers, like traders, are scrambling to stay ahead of the curve. Here are 10 of their recent observations.

There's more to this story

Find more charts, plus interactive elements and citations for research mentioned, at **chicagobooth.edu/review**.

1. LLMs can be trained specifically for finance and other domains.

Modern LLMs have significantly advanced the capabilities of natural language processing, essentially learning from giant data sets that represent a large swath of human knowledge. But some research indicates it may be possible to create more specialized, domain-specific LLMs that, at times, outperform the general-purpose models such as GPT-4.

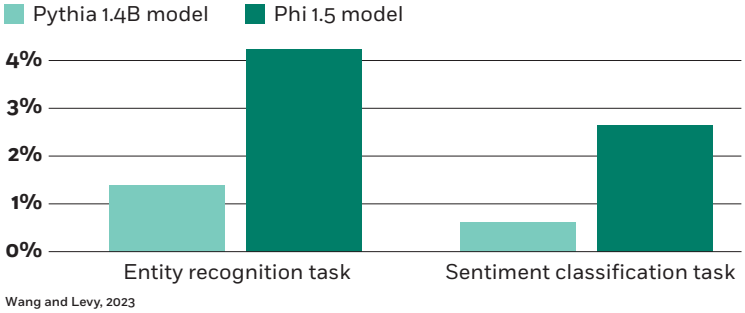
Fine-tuning a smaller model has benefits beyond just customization for a particular task. It also lowers the computing costs, improves data privacy, and produces a tool that runs much faster than general-purpose models—possibly even on mobile devices.

Motivated by this idea, Chicago Booth research professional Siyan Wang and Booth's Bradford Levy created a finance-focused data set, called BeanCounter, which contains over 159 billion tokens extracted from corporate disclosures filed with the Securities and Exchange Commission. (A *token* is a word or part of a word.) For reference, OpenAI has disclosed that it trained GPT-3 using 300 billion tokens, primarily from Common Crawl, a nonprofit repository of online data. The researchers note that BeanCounter contains less than 0.1 percent of the data in Common Crawl-based data sets. What's more, they examined content directed toward various demographic identities and find that the content in BeanCounter tends to be substantially less offensive or harmful and more factually accurate.

Could a smaller data set ever produce an LLM whose performance could match that of GPT-4 or a similarly broad model? Wang and Levy say they have evidence that their LLM trained on BeanCounter actually does better. They used it to continuously pretrain two existing small, open-source LLMs. In finance-related tasks including sentiment analysis, the models pretrained on BeanCounter showed performance improvements over their base models. Both models also registered an 18-33 percent reduction in the level of toxic text generated after being updated with the data set.

Data quality matters a lot, says Levy, arguing that an LLM trained on fewer data points can perform well if they're high quality. The findings highlight the capabilities of smaller LLMs that are customized for various tasks or domains—and that work faster and cost less than large, generalized models.

Improvements in performance of finance-related tasks with additional training on BeanCounter data



Wang and Levy, 2023

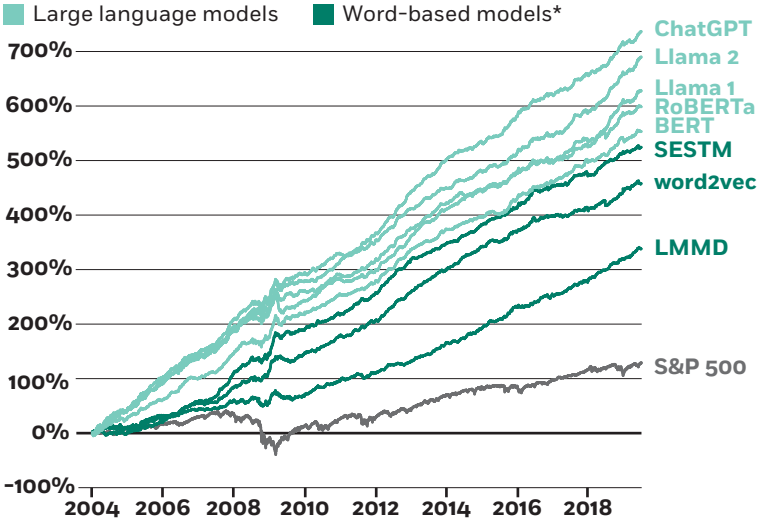
2. LLMs can improve stock return predictions.

When it comes to analyzing the sentiment of news articles, LLMs are far better than other models that came before them, research suggests. Booth PhD student Yifei Chen, Yale's Bryan T. Kelly, and Booth's Dacheng Xiu applied LLMs and traditional word-based models such as Google's word2vec to analyze the sentiment of business-news articles and their headlines in 16 global equity markets. And when they used the sentiment scores to make stock return predictions, the portfolio informed by LLMs outperformed those of word-based models. Thanks to their nuanced grasp of the meaning and structure of language, LLMs demonstrated a better comprehensive understanding of what was being said, and this led to a deeper interpretation of the news and greater predictive accuracy.

Models for predicting stock returns typically rely on variables focused on a company's characteristics, financial data, and historical returns. By creating a news sentiment variable and adding it to a predictive model, Chen, Kelly, and Xiu introduced an alternative data source, which also provided an opportunity for the model to capture additional data. For example, any information released overnight was missed by past return variables but contained in the sentiment variable.

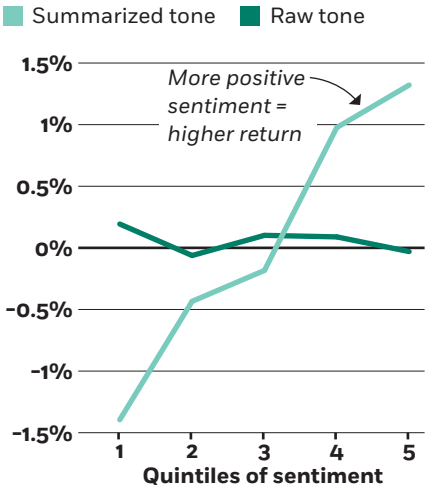
Their research reveals a pronounced short-term momentum effect linked to news and suggests LLMs may offer promising opportunities for investors wanting to capture news sentiment in their models. Their simulations for larger LLMs such as RoBERTa (similar to one of the best-known LLMs, BERT–bidirectional encoder representations from transformers–but trained on a larger and more diverse data set) and Llama 1 and Llama 2 by Meta (similar to OpenAI's GPT-based models) achieved exceptional risk-adjusted returns. They saw Sharpe ratios above 4, a level that proprietary trading funds eagerly seek.

Cumulative returns of portfolios sorted by sentiment scores from different models



*SESTM = Sentiment extraction via screening and topic modeling; LMMD = *Loughran-McDonald Master Dictionary*
Chen et al., 2023

Cumulative abnormal returns over a two-day period



Kim et al., February 2024

3. LLMs can produce useful summaries.

Companies disclose a lot of unstructured textual information in annual reports, with the management discussion and analysis sections found in 10-K filings being a salient example. ChatGPT can quickly distill the gist of what's being shared by summarizing both the MD&A and earnings call transcripts, research demonstrates.

In a study, Booth researchers Alex Kim, a PhD student; Maximilian Muhn; and Valeri Nikolaev demonstrate how using GPT-3.5 Turbo can enhance clarity by stripping away boilerplate language, generic phrases, and less relevant details, offering a more accurate reflection of investor-relevant sentiment contained in complex corporate disclosures.

The researchers find that the sentiment of the GPT-3.5 Turbo-based summaries of the earnings announcements and 10-K filings, as opposed to the sentiment of the raw text, better explained the contemporaneous abnormal returns that resulted from investors reacting to these events. This suggests an opportunity for investors to use LLM summaries to enhance signals for trading around earnings calls and the release of 10-Ks. (For more, read "ChatGPT could help investors make more informed decisions," in the Fall 2023 issue and online.) Indeed, over the past year, several AI startups have emerged that generate summaries and allow customers to query corporate filings and communications.

4. LLMs can identify corporate risk.

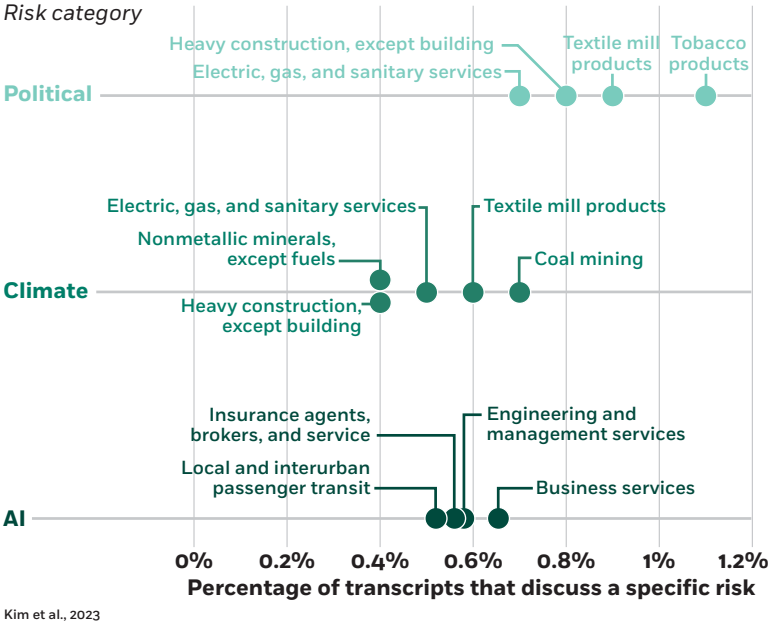
Some information is relatively straightforward to extract from corporate filings and earnings calls, but other information is trickier—such as certain types of risks facing a company. After all, managers on a public earnings call aren't usually keen to highlight their challenges.

However, in another paper, Kim, Muhn, and Nikolaev demonstrate the potential of LLMs to detect hard-to-quantify corporate risks, even when those risks are only indirectly disclosed. Their research suggests that an LLM can make inferences by using the vast amount of information on which it is trained to pick up on complex and nuanced relationships between statements scattered throughout a transcript's text.

The research focused in particular on political, climate, and AI risks affecting corporations. In the past, a number of different researchers have tried to pull insights on corporate risk from earnings call transcripts using natural language processing. They've had somewhat limited success due to executives' careful language choices and the algorithms' inability to understand the deeper context of what's being discussed. For example, a call transcript may contain a risk-related discussion without explicitly mentioning risks anywhere.

But GPT-3.5 Turbo connected statements made throughout a transcript and leveraged its vast knowledge base to infer, or read between the lines to discover, the risks, find Kim, Muhn, and Nikolaev. As a result, they say, the risk measures GPT-3.5 Turbo produced were capable of more accurately predicting volatility in a company's stock price following its earnings call. The LLM was even able to capture newly emerging risks that were not commonly seen in its training data, including risks associated with AI itself. (For more, read "AI reads between the lines to discover corporate risk," Spring 2024 and online.)

Industries with the highest LLM-based risk assessment scores



Kim et al., 2023

5. LLMs can find new information in disclosures.

In a different spin on making sense of lengthy, dense corporate disclosures, Booth's Anna Costello, Levy, and Nikolaev developed LLMs that can spot new information. Rather than summarize documents, their method retains the original content and instead highlights the portions that are likely to be surprising.

The researchers first built a base LLM pretrained on financial disclosures from various companies, text that altogether totaled more than 35 billion tokens. They then created a firm-specific LLM for each company by further training the base model on that company's past regulatory filings. By using only contemporaneous data, the researchers made sure their measure learned solely what investors could have known about a company at the time.

Information theory holds that surprising events are those that investors assign a relatively low probability of happening. Along these lines, LLMs work by modeling the probabilities around the next word (or partial word) in a sequence of text. The researchers precisely measured the level of surprise associated with each word in a filing relative to the content on which the LLM had been trained.

They then used the notion that prices should fully reflect all publicly available information to validate their measure, finding that it explains a large portion of the short-term market reaction to corporate filings and is predictive of future returns. This future predictivity is small, however—consistent with market efficiency and limits to arbitrage, they write.

While this work is focused on corporate disclosures, the researchers say that their method is general enough to be applied to other settings such as supply-chain contracts and legal documents, or even other modalities such as images and video. Given the novelty of the method, notes Levy, the University of Chicago is pursuing a patent on the technology.

6. LLMs can predict earnings more accurately than humans.

Financial-statement analysis requires quantitative skills, logical reasoning, critical thinking, and complex decision-making–so one might think that it's a domain in which humans still have a leg up on LLMs.

Research suggests this may soon change. Kim, Muhn, and Nikolaev find that LLMs, and specifically GPT-4 Turbo, can simulate professional financial-statement analysis, and in a way that outperforms humans.

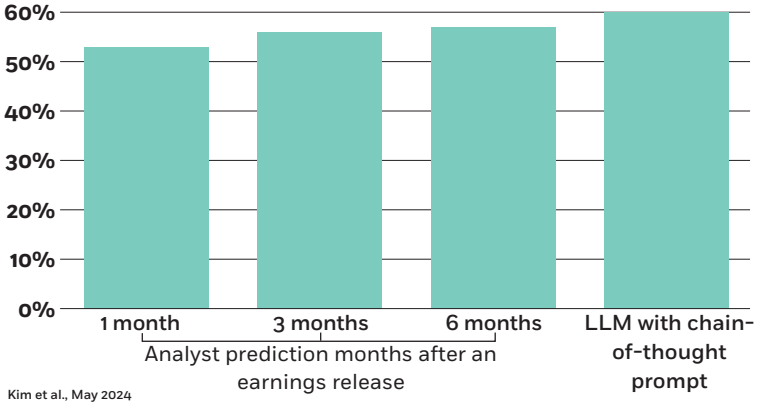
The researchers provided GPT-4 Turbo with anonymous balance sheets and income statements with the management discussion and analysis sections removed so that the LLM did not have any textual clues. They standardized the statements by making sure all labels matched a template and replacing dates with *T*, *T-1*, and the like.

Then they used a "chain of thought" prompt to instruct the LLM to solve problems step-by-step and with reasoning as a human would. They wanted it to use the same thought process as an analyst in noting trends in the statements and computing key financial ratios. While earlier LLMs including GPT-3.5 are notoriously bad at math, the newer GPT-4 Turbo leveraged its understanding of math concepts and combined its computations with economic reasoning to deliver insights about companies, the researchers find.

Its predictions outperformed the consensus forecasts of professional financial analysts (60 percent accuracy versus 53 percent). Furthermore, its accuracy was on par with a sophisticated machine-learning model specifically trained to predict the direction of earnings.

The paper suggests LLMs have a relative advantage over analysts, who in some instances may struggle to come up with an accurate forecast (and hence issue differing forecasts) or display bias. In further study, the researchers find that human analysts' forecasts complemented GPT-4 Turbo's forecasts, indicating that professionals still provide valuable insights about companies and markets that aren't reflected in financial statements. They conclude that LLMs have the potential to play a central role in financial decision-making–complementing humans rather than replacing them.

LLM's prediction accuracy versus human analysts



Kim et al., May 2024

THE EVOLUTION OF AI IN FINANCE

To appreciate the edge that artificial intelligence can bring to the financial markets, it's worth understanding how fast the technological landscape has changed for investors. It has been propelled by research that has incorporated advanced techniques from AI, particularly from several subfields that have played a crucial role.

One is machine learning, which involves training algorithms to learn patterns and make predictions from data. The term dates back to 1959, but the area of study began to receive a lot more attention starting in the early 2000s as computational power increased and the internet helped support a trove of data available to train ML models.

In the past five years, researchers have embraced ML to solve finance problems. In 2020, Booth PhD student Shihao Gu, Yale's Bryan T. Kelly, and Booth's Dacheng Xiu summarized the performance of diverse ML models when applied to finance. They presented various models predicting stock returns and compared them in terms of efficiency and accuracy. The best performers were trees and neural networks—statistical methods modeled on decisions and outcomes, and on the human brain, respectively. The paper has been widely cited in research, racking up more than 1,800 citations so far.

That same year, City University of Hong Kong's Guanhao Feng, Yale's Stefano Giglio, and Booth's Xiu created an ML method to evaluate factors and identify those most relevant for asset prices. In 2021, Booth's Stefan Nagel published a book, *Machine Learning in Asset Pricing*, to explain how ML tools, which were not originally developed for finance, could be applied to empirical research in pricing and theoretical modeling of financial markets.

Researchers have since used ML to predict prices and construct portfolios, among other tasks.

Meanwhile, finance research has progressed in the subfield of natural language processing, an area in which ML techniques are turned on language itself to mine information from text. Early adopters of language tools included Shanghai Jiao Tong University's Feng Li (a graduate of Booth's PhD program), who in 2008 studied the relationship between the readability of 10-K filings and corporate performance. He found that companies with longer and more difficult-to-read reports tended to have poorer earnings.

In separate but related work, University of Notre Dame's Tim Loughran and Bill McDonald explored sentiment analysis of 10-Ks, finding in a 2011 paper that the existing dictionary of words used to determine the sentiment of a text was not well suited to the financial domain. For example, words such as *liability*, *cost*, and *tax* were scored as negative for sentiment using the traditional dictionary, but these words are not necessarily negative when used in a financial context. Loughran and McDonald in turn created a dictionary tailored to finance.

Other researchers have developed new techniques for analyzing textual data. Boston University's Tarek Alexander Hassan, Tilburg University's Stephan Hollander, Frankfurt School of Finance and Management's Laurence van Lent, and London Business School's Ahmed Tahoun (then a research scholar at Booth) published research in 2019 that used a simple algorithm for assessing political risk in earnings call transcripts. It counted bigrams (two-word combinations including *the constitution* or *public opinion*) used in conjunction with the words *risk* and *uncertainty*, or their synonyms, to identify potential risks to companies. The higher the count, the greater the political risk for the company, the research finds. Subsequent papers resulted in a startup, NL Analytics, that works with central banks and international organizations to use these methods for economic surveillance.

The jumps that led to deeper understanding

Finance and accounting have long sought to learn from text. Economists have, too, and originally used a "bag-of-words" model. This relies on counting word frequency in a text—for example, how many times does a document include the words *capital* and *spending*? In this case, the more frequently these words occur, the more likely it is that the document discusses corporate policies.

This method is straightforward: in 1963, the late Frederick Mosteller and the late David L. Wallace used it to argue that James Madison, not Alexander Hamilton, had written 12 of the 85 essays and articles in the *Federalist Papers* whose authorship had been in dispute. By counting commonly used words in Madison's and Hamilton's known texts, they could compare them with the count of those words in the disputed articles in the *Federalist Papers*.

The method is also limited, however. It doesn't take into account potentially important information such as grammar or the order in which words appear. As a result, it's unable to capture much in terms of a document's context. A company's 10-K filing might report that "Increased transportation costs have offset our revenue gains," and bag-of-words may interpret this as a positive statement—after all, the word *increased* and the phrase *revenue gains* might seem confident. But it misses the fact that *increased* taken with *costs* is negative and that *offset* changes the meaning of *revenue gains*.

Researchers at Google took a big step toward incorporating this context in 2013 when the company introduced word2vec, a neural network–based model that learns vector representations of words and captures the semantic relationships between them. Vectorization enabled ML models to process and understand text in a more meaningful way. If you have three related words, such as *man*, *king*, and *woman*, word2vec can find the next word most likely to fit in this grouping, *queen*, by measuring the distance between the vectors assigned to each word.

And in a 2017 paper, a team of researchers led by Ashish Vaswani, who was then at Google Brain, introduced what's known by practitioners of deep learning as transformer architecture. Transformers form the basis of the large language models we know today and represent a significant improvement over previous architectures in their ability to understand and generate human language, which word-based models could not do.

One prominent LLM, BERT (bidirectional encoder representations from transformers), is used to understand the context of words but was not designed to generate text. It works by considering the words that appear before and after a particular word to decipher its meaning.

Meanwhile, GPT (generative pretrained transformer) is able to predict the most likely next word in a sequence based on the text leading up to it. For example, finish this sentence: "Why did the chicken cross the _____?" Your brain automatically fills in the blank with the word *road* as the most probable next word, even though many other words would work here, including *street*, *highway*, or maybe even *yard*. GPT does the same thing. Its parameters can be set, however, so that it doesn't always choose the highest-probability word. This allows more creativity in the text it generates.

Now these LLMs, too, are tools that are being applied to finance, enabling researchers and practitioners in the field to extract increasingly valuable insights from data of all kinds.

Go to **chicagobooth.edu/review** to see citations for research mentioned in this article.

7. LLMs can signal stock crashes.

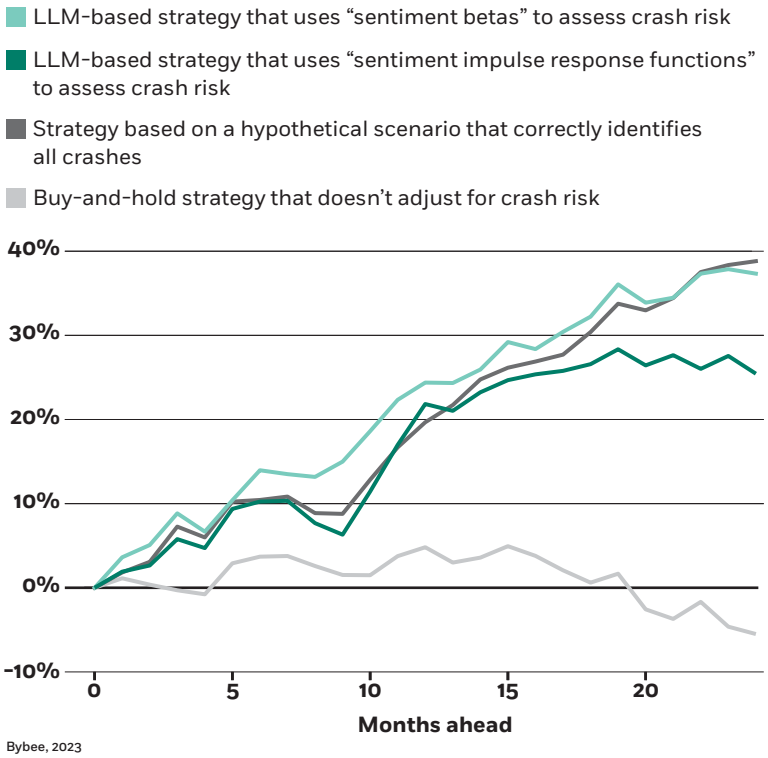
LLMs can predict key financial and economic indicators, finds a study by Booth's Leland Bybee. His research creates a method for doing so by applying an LLM to news articles and then forecasting financial and economic measures such as the S&P 500 and the Consumer Price Index.

Applying the method to 100 years of news articles, Bybee produced a time series of economic beliefs. The predictions made by the LLM aligned closely with those recorded in investor and CFO surveys, as well as with equity fund flows.

And when he used the method to investigate behavior during financial bubbles, he finds that the more sentiment (rather than fundamentals) fueled a rise in an industry's stocks, the higher the probability of a crash and lower future returns. This suggests that sentiment-driven mispricing can predict bubbles.

Bybee tested these findings with an estimated trading strategy that held portfolios of stocks that were all from the same industry, were generally rising, and were predicted not to crash. That prediction was based on a cutoff threshold in the sentiment measure the researcher produced. This LLM strategy successfully avoided 80 percent of the industries that had sharp downturns and significantly outperformed a similar but more naïve strategy.

Cumulative returns on a portfolio of industry stocks with low crash risk



Bybee, 2023

8. LLMs can see through boilerplate tax disclosures.

Tax audits are hugely important for companies and their investors, but you wouldn't know it from the boilerplate tax disclosures companies voluntarily make in corporate filings. Is an audit imminent? Has it just concluded? Are authorities about to levy a fine or challenge a company's tax-planning strategies? This has all been difficult for investors to figure out.

However, LLMs can make sense of the hard-to-parse disclosures and extract useful signals for investors, suggests research by City University of London's Ga-Young Choi and Booth PhD student Kim.

The researchers used GPT-4 to analyze about 20,000 10-K filings from 2010 to 2021, extract each company's relevant tax and audit information, and track the changes in language from one year to the next.

These differences may indicate some potential corporate risk, according to the study. Active tax audits effectively deterred tax avoidance but led to increased stock volatility and reduced capital spending, the researchers find. Even after an audit had concluded, companies in their sample tended to continue to decrease tax avoidance strategies, capital investments, and new debt issuance.

The research suggests that an LLM can be applied to corporate disclosures to tease out a company's current audit status and help anticipate and avoid any potential related fallout.

9. LLMs can measure and identify corporate policies.

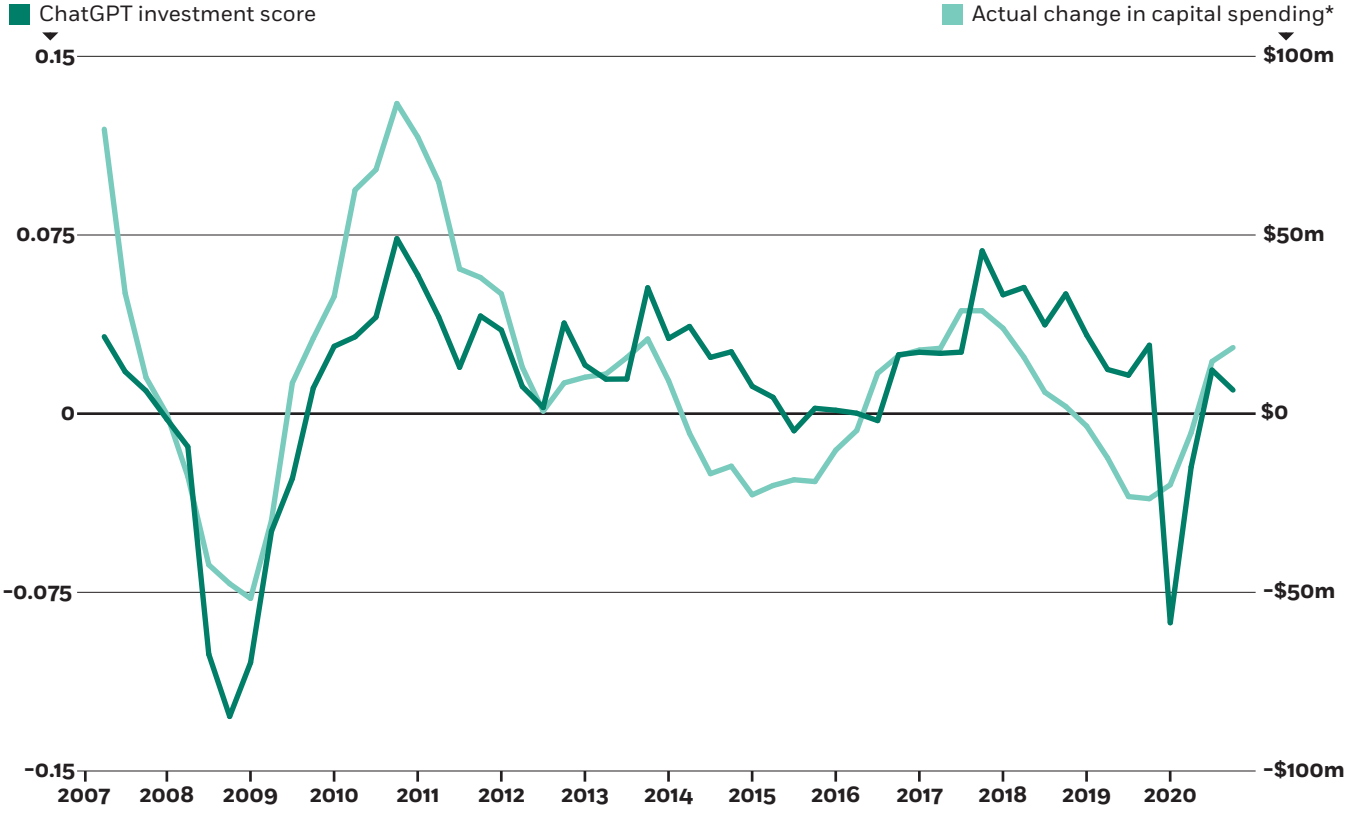
On earnings calls, investment policies often aren't stated simply, or even at all. For example, an executive may say, "We are investing in growth initiatives." While the line doesn't directly state as much, this might imply some large, upcoming capital expenditures that could affect near-term profitability.

Researchers from Georgia State, Manish Jha, PhD student Jialin Qian, and Baozhong Yang, along with Booth's Michael Weber, designed a method using ChatGPT that they suggest is capable of discovering sometimes-hidden policies. The underlying LLM can analyze call transcripts and predict future corporate policy changes—such as shifts in capital investments, dividend levels, or head count, their research finds.

The researchers used ChatGPT to generate a likelihood score for changes in corporate policies. The score was validated by its alignment with CFO survey responses about corporate investment plans. Their method's predictions for capital spending and the actual capital expenditures were highly correlated.

ChatGPT was able to decipher the corporate policy changes from the transcripts with a high degree of accuracy, the researchers write. The scoring system they devised could serve as a tool for investors by revealing potential corporate policy shifts not fully priced into the market. In the research, high investment scores were linked to notable negative abnormal returns over subsequent quarters, suggesting this tool can offer an advantage in portfolio management, especially in conjunction with other analyses such as Tobin's Q, which is used by investors to evaluate corporate policies. (For more, read "AI can discover corporate policy changes in earnings calls," online.)

ChatGPT investment score compared with realized investment from a sample of companies



*Difference in companies' average capital spending four quarters after and four quarters before the current quarter

Jha et al., 2023

IMAGES AND AUDIO ARE NOW DATA TOO

The data-analysis revolution that turned words into analyzable data continues to progress. Now models are turning images, audio, and visual files into data as well. Large language models can capture meaning from text data in a way that wasn't possible before, and these media are mostly untapped territory.

Images

Price charts—the actual charts, not the data underlying them—have been used to predict stock returns.

By applying a deep-learning algorithm called a convolutional neural network to analyze images of historical stock charts, University of Chicago PhD student Jingwen Jiang, Yale's Bryan T. Kelly, and Chicago Booth's Dacheng Xiu extracted predictive patterns and converted them into trading signals. The image patterns achieved more accurate return predictions than common trend signals used in technical analysis, they find.

Meanwhile, other research finds that financial analysts' facial characteristics are associated with their forecast outcomes. Using artificial intelligence and machine-learning models, Baruch College's Lin Peng, University of California at Los Angeles' Siew Hong Teoh, Chinese University of Hong Kong's Yakun Wang, and Cornell PhD student Jiawen Yan scored the LinkedIn photos of approximately 800 sell-side stock analysts for characteristics such as trustworthiness, attractiveness, and dominance and examined their relation with the analysts' earnings forecast accuracy over the past three decades.

Their study finds that analysts who scored high on trustworthiness produced more accurate forecasts. The researchers surmise this is likely because people are more comfortable sharing information with individuals they trust, thus these trustworthy-looking analysts gained more information that improved their forecasts.

However, their findings also reveal a striking gender disparity. Male analysts with high dominance scores made more accurate forecasts compared with those with lower scores. Conversely, female analysts perceived as dominant had less accurate forecasts than those with lower dominance scores. Meanwhile, a higher dominance score significantly increased the likelihood that a male analyst would be voted an All-Star, a marker of professional prestige—yet this score substantially decreased the chances for a woman, despite female analysts having, on average, more accurate forecasts than their male counterparts.

The researchers interpret the findings as potential evidence of gender discrimination in this labor market, arguing that the perception of increased masculinity contradicts the female stereotype, making female analysts less likable. (For more stories and videos about what else researchers are learning from faces, go to **chicagobooth.edu/review**.)

Audio files

Managerial vocal delivery quality is associated with real-time market reactions during earnings calls, research suggests. When investors strain to understand what is said in the call due to mumbling, mispronunciation, or just plain lazy diction, the market reaction to the earnings call tends to be more subdued, find Seoul National University's Bok Baik, Booth PhD student Alex G. Kim, MIT PhD student David Sunghyo Kim, and Artificial Society's Sangwon Yoon. The researchers used a DL algorithm to convert audio files from the earnings calls into letters, which were then combined into words and ultimately text. (For more, read "On earnings calls, do executives mumble on purpose?" Summer 2024 and online.)

In another study, researchers from Ruhr University Bochum—Jonas Ewertz, Charlotte Knickreh, Martin Nienhaus, and Doron Reichmann—used vocal cues to predict the future earnings of a company. They first visualized the vocal cues of managers on earnings calls with a mel spectrogram, which converts frequencies to the mel scale, a measure that represents sound in a way that humans typically hear it. They fed those images into a DL algorithm. Their model's predictions for changes in future earnings significantly outperformed models that used numerical and text data.

Similarly, UC Berkeley's Yuriy Gorodnichenko, University of York's Tho Pham, and University of Birmingham's Oleksandr Talavera analyzed the influence of vocal emotion on financial variables such as share price, volatility indices, interest-rate risk, inflation expectations, and exchange rates. Using a DL model to detect vocal emotions in the press conferences after Federal Open Market Committee meetings, the researchers find that a significantly positive tone led to higher share prices. In fact, they write, "switching the tone of the press conference from negative (-1) to positive (+1) could raise S&P 500 returns by approximately 200 basis points."

Videos

With the advantage of both images and audio, videos may reveal information that can't be uncovered by other media on their own, research suggests. University of Washington's Elizabeth Blankespoor, Hong Kong Polytechnic University's Mingming Ji, University of Hong Kong's Jeffrey Ng, and PolyU's Jingran Zhao created a sample of about 500 CEO earnings-announcement-related interviews broadcast on CNBC from 2013 to 2017. They find that when CEOs' facial expressions were incongruent with their earnings news, the dispersion across analysts' forecasts increased.

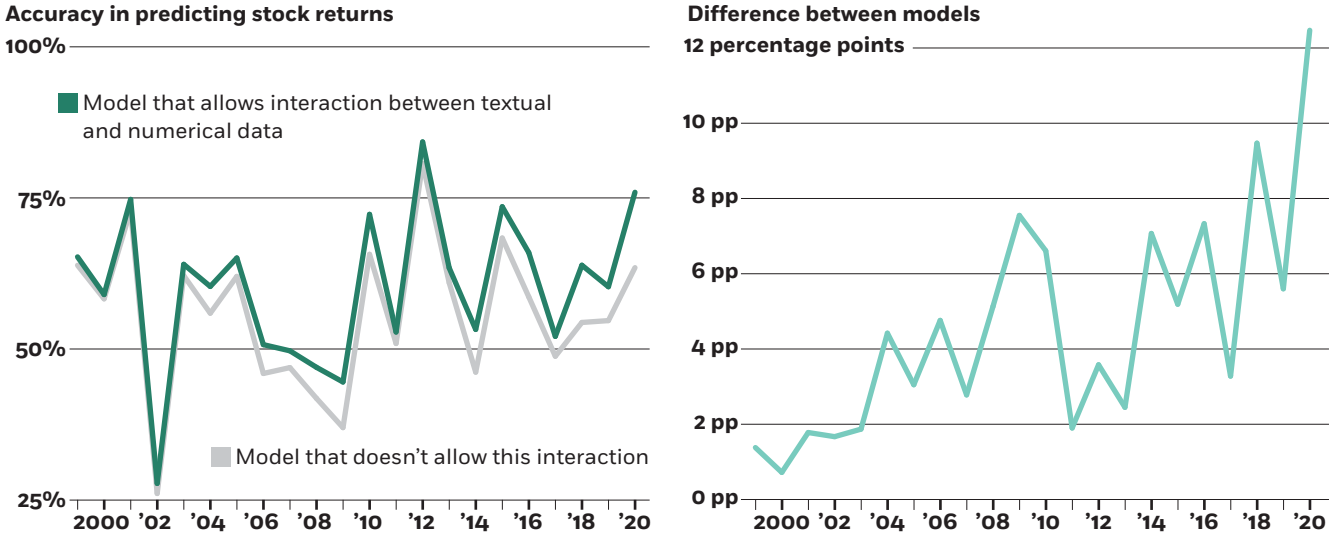
Go to **chicagobooth.edu/review** to see citations for research mentioned in this article.

10. LLMs can combine insights from text with numbers to produce more accurate predictions.

LLMs can incorporate the textual information in the MD&A section of a 10-K filing to enhance the value of the numerical information disclosed by a company, suggests research by Kim and Nikolaev. They used BERT to contextualize accounting numbers by incorporating textual information and find that this improved the accuracy of predictions about future earnings, cash flows, and stock returns.

Specifically, integrating textual information about demand trends and strategic plans for a company with the numerical data about profitability improved the model's performance compared with using solely numerical or textual data, according to one of two related papers they wrote on the topic. Also, predictions of share prices and portfolio performance improved when the researchers included in their model a measure that they created, context-adjusted profitability.

The findings suggest that investors can improve their strategies by using LLMs to incorporate these textual data. (For more, read "Large language models can improve stock market forecasts," Spring 2024 and online.)



Kim and Nikolaev, 2023

The takeaway: Make use of LLMs

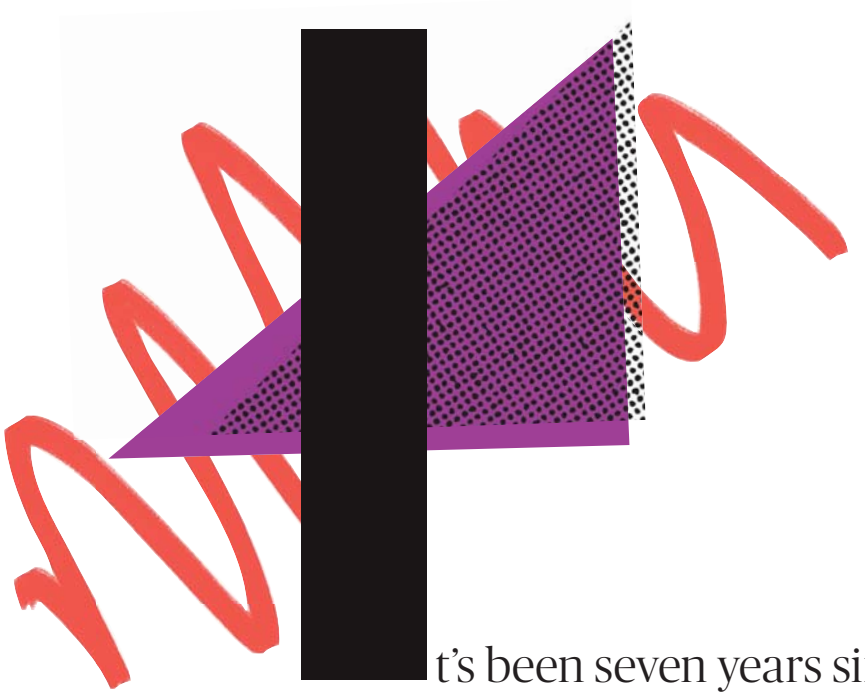
AI is transforming practically every sector of the economy. The technological progress being made has implications for everything from job recruiting to medical diagnosis to filmmaking. (For more, read "AI is going to disrupt the labor market. It doesn't have to destroy it," Winter 2023 and online.) In finance, LLMs are mining public data to find varied and largely unexploited investment opportunities—and are evolving from being analytical tools to capable decision-makers, paired with investors in the ongoing hunt for profit.–CBR

THE CASE FOR A RETRO TAX CODE

Should US lawmakers design the future to look like 1997?

BY NEIL WEINBERG
ILLUSTRATIONS BY NATE KITSCH


Nintendo GAME BOY



It's been seven years since US president Donald Trump took to a Washington, DC, podium to sing the praises of a tax bill, soon to become the Tax Cuts and Jobs Act of 2017—better known to many as the "Trump tax cuts."

"My administration is working every day to lift the burdens on our companies and on our workers so that you can thrive, compete, and grow," Trump told assembled members of the National Association of Manufacturers. "And at the very center of that plan is a giant, beautiful, massive—the biggest ever in our country—tax cut." The TCJA, which passed along a near-party-line vote and was signed that December, delivered the largest corporate rate cut as a percentage of gross national product in US history. A *New York Times* article called it "the most sweeping tax overhaul in decades." It was also, according to *FiveThirtyEight*, "one of the least popular tax plans since Ronald Reagan's day," supported by about one-third of voters.

Nearly a decade on, taxes remain a political lightning rod. Critics say the TCJA provided a windfall to the wealthy and sent the federal deficit soaring. "The Trump tax cuts kicked working families to the curb, while it added about $2 trillion to the national debt, all so that Republicans could give a massive handout to their wealthy pals and donors," said Senator Elizabeth Warren (Democrat of Massachusetts) last November at a hearing of the Senate Finance Committee. On the other side, former senator Phil Gramm (Republican of Texas) told the House Ways and Means Committee that the TCJA "created an environment in which people invested more money and created more jobs."

Americans should steel themselves for more such debate. "The spotlight may be on the US 2024 presidential election, but tax executives can't afford to take their eyes off 2025 and beyond," PwC's Ken Kuykendall wrote recently on the audit firm's website. Several provisions of the TCJA are set to expire at the end of 2025, setting up what Representative Blake Moore (Republican of Utah), a member of the Ways and Means Committee, is calling a "Super Bowl of tax" year.

As with so many issues in the United States, the choice of which way the country should head on taxes seems stark. On one side is a vision of continued tax cuts to fuel rising economic growth. On the other is a tax regime that wants corporations and wealthy households to pay more.

But policymakers on both sides of the aisle acknowledge a need to balance the federal budget, a dance between spending and revenue. Corporate tax collections fell below $250 billion annually after the TCJA became law, down from about $375 billion a year during the middle of the previous decade, point out Princeton's Owen Zidar and Chicago Booth's Eric Zwick. By their calculations, the decline in corporate tax revenue was larger than the increase in business investment spurred by the cuts.

Tax reform doesn't rank near the top of many voters' priorities. But policy wonks of all stripes are well aware that the outcome of the presidential and congressional races will determine who has the upper hand in the coming negotiations when the existing provisions expire.

Could any reform garner bipartisan support? Zidar and Zwick argue they have a proposal that could do exactly that. It would raise more than $4 trillion in additional revenue over a decade, and would do so by tweaking the tax code rather than overhauling it. And it would look a lot like the tax code did in January 1997.

The researchers analyzed how the TCJA's business tax provisions have performed, with an eye to proposing adjustments that would raise the revenue needed to tame run-away deficits and fund priorities such as defense and health programs—but without choking the golden goose that is the world's largest economy.

They treated the current system like a panel of dials: one for taxes on corporations, another for dividends, a third for estates, and so on. They then analyzed how much revenue could be raised by turning up the tax-rate dial in specific areas and estimated the likely effects on economic activity.

How the landscape has changed

To assess the merits of Zidar and Zwick's back-to-the future proposal, consider how dramatically legislators have overhauled the federal government's revenue-generation machine in recent decades. The TCJA revised taxes on foreign income, exempting it from the newly lowered corporate rate of 21 percent and replacing it with a worldwide system under which foreign profits are taxed only when they're repatriated. However, arguably the most notable feature of the law was a series of business tax cuts that have changed how individuals structure their enterprises and categorize the income they earn through them.

Many Americans are familiar with C corporations, the *corporate* in "corporate America." Lesser appreciated are pass-through businesses such as S corporations, partnerships, and sole proprietorships, in which profits and losses pass through to individual owners, who pay according to personal income tax rates.

C corps were the dominant business structure until the Tax Reform Act of 1986 cut individual rates and led taxpayers to shift a significant portion of business activity to minimize taxes. During the succeeding decade, pass-throughs eclipsed C corps as the primary generators of US business income, and partnership structures led the way in the most recent years, the researchers explain. "Much of the charted rise of pass-through income reflects simple recategorization: to take advantage of lower tax rates, business owners have reclassified C-corporation income as pass-through," write Zidar and Zwick in an essay for the Aspen Institute.

These business owners are, for the most part, among the highest earners in the US workforce, according to research by the Treasury Department's Matthew Smith, University of California at Berkeley's Danny Yagan, Zidar, and Zwick—often doctors,

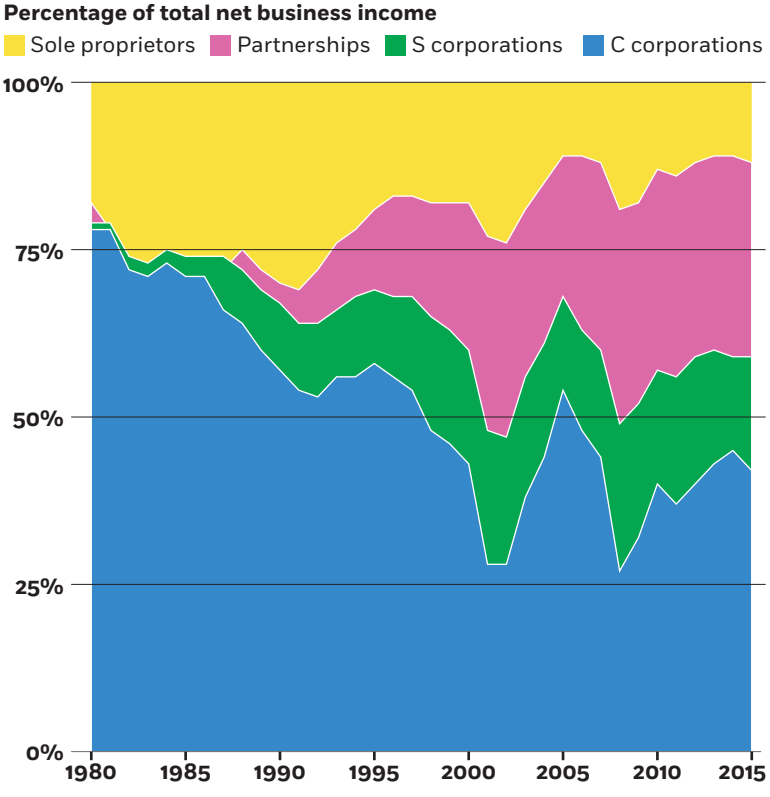
The Tax Cuts and Jobs Act of 2017 dramatically cut C corporations' federal tax burden.

lawyers, or owners of middle-market businesses such as car dealerships and drink distributors. Roughly two-thirds of every dollar currently earned by pass-throughs accrues to individuals who are already among the top 1 percent of earners, write Smith, Yagan, Zidar, and Zwick. More than 1 million business owners, each making at least $390,000 annually, reported some pass-through income as of 2014, and more than 140,000 people making at least $1.6 million annually did so. These income thresholds are likely 50-70 percent higher today, estimates Zwick. As for the number of the top earners reporting pass-through income, he and Zidar write that it far surpasses the number of executives at public companies, "who have been the focus of much public commentary about inequality."

The TCJA tipped the scales back in favor of traditional C corps by granting them larger rate cuts than pass-throughs. The law dramatically cut C corporations' federal tax burden. Listed companies had their effective tax rates cut by 9 percentage points, Zidar and Zwick calculate. C corps saw their headline tax rates slashed from 35 percent to 21 percent. The alternative minimum tax was abolished. Among other provisions, businesses were allowed to immediately expense many equipment investments. In some cases, the tax code grants taxpayers so many deductions that they may pay little or no tax.

The rise of pass-through businesses

The share of business activity accounted for by pass-through entities has increased since the 1986 Tax Reform Act lowered the top individual tax rate.



Zidar and Zwick, 2023

But the high earners who control many pass-throughs have also been big beneficiaries of a TCJA provision that cut the top marginal tax rate on personal income from 39.6 percent to 37 percent. Workers lower down the pay spectrum or in certain sectors enjoyed an even larger cut in their top marginal tax rate, from 37 percent to 29.6 percent. Many high earners still regard the pass-through structure as the more attractive option, write the researchers. And whatever business structure these owners favor, the TCJA probably delivered them a tax cut.

Grading Trump-era tax reform

The TCJA's backers predicted the law would pay for itself, boost domestic investment, and benefit rank-and-file workers. The law's authors took several steps to prevent it from becoming a budget buster. For example, they reduced deductions for business losses and interest expenses. They trimmed tax credits for companies that incur research and development costs. And they eliminated favorable rates previously available to domestic manufacturers, among others, under the Domestic Production Activities Deduction.

It will take time to establish how those predictions have played out—and the picture is complicated by other factors that affected the economy in the wake of the TCJA, including the COVID pandemic and the trade war with China. Still, Zidar and Zwick are among the researchers analyzing data to paint a picture of the law's effects.

Some predictions may have been too rosy, starting with those relating to corporate tax collections. The amount of revenue the government brought in from corporate taxes fell from about 1.8 percent of GDP to as little as 1 percent, find Zidar and Zwick. The 2022 figure of 1.3 percent was still well below pre-TCJA levels, they write.

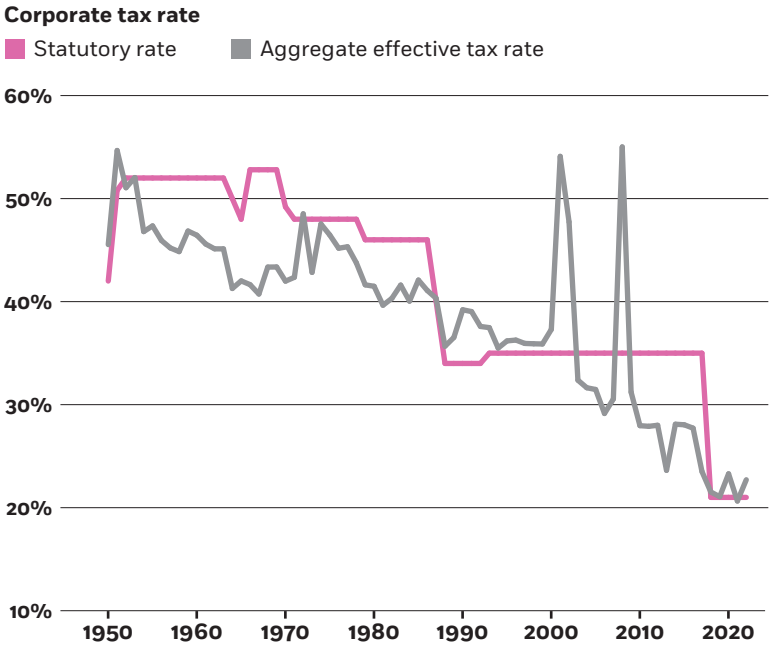
Investment hasn't made up for lower collections, according to Harvard's Gabriel Chodorow-Reich, Smith, Zidar, and Zwick. In 2018, the president's Council of Economic Advisers forecast that the TCJA would increase domestic investment in equipment and structures by 9 percent, or roughly $300 billion, and several studies indicate that C corps, including publicly listed enterprises, did substantially increase investments. Chodorow-Reich, Smith, Zidar, and Zwick find that C corps with the mean tax change, relative to those whose tax burden didn't change, increased their domestic investment by 20 percent. But the law led corporate tax collections to fall 41 percent, and investment helped offset that by just 2 percentage points on average over 10 years, according to their calculations.

In the first quarter of 2018, 95 companies in the S&P 500 said on earnings calls that they planned to increase investment because of the TCJA, other research finds. According to an analysis by MIT's Michelle Hanlon, University of North Carolina's Jeffrey L. Hoopes, and University of Michigan's Joel Slemrod, companies that expected to save the most because of the tax cuts were also most likely to announce higher investment and benefits for workers. "The business provisions of the TCJA arguably made investment more attractive by reducing the tax-adjusted cost of capital," the researchers write.

But they also find that companies whose political action committees had donated more to Republican than Democratic candidates were likelier to announce worker benefits, which they point to as an indication that political factors, not just economic ones, may well have been at work. Companies may have sought to broadcast the news that tax cuts had widespread benefits, perhaps to fend off any

A win for C corporations

The Tax Cuts and Jobs Act of 2017 lowered the corporate tax rate from 35 percent to 21 percent, tipping the scales back in favor of traditional C corps.



Zidar and Zwick, 2023

potential reversals. "We also find that a corporation is marginally less likely to announce a TCJA-tied worker benefit if it is headquartered in a red state," write Hanlon, Hoopes, and Slemrod. The benefits to workers were tracked and publicized by the Americans for Tax Reform, a right-leaning advocacy group, the researchers report.

In arguing that the Trump administration's estimate may have overstated investment gains, Zidar and Zwick point to research—from University of California at Los Angeles' Patrick Kennedy, the Federal Reserve Board's Christine Dobridge, and the Joint Committee on Taxation's Paul Landefeld and Jacob Mortenson—that used tax data to compare how similarly sized C corps and S corps responded to the cuts. C corps saw the bigger cuts, and increased investment more. On the basis of this finding, Zidar and Zwick estimate that midsize C corps invested up to $81 billion more than before. However, the tax cut removed $88 billion in corporate tax revenue from the coffers, which far exceeds the implied tax revenue gains from the higher investment for these firms.

The law may have likewise defied assertions that it would lift all boats. In 2017, administration economists predicted that the tax reform would increase average household income by at least $4,000 annually as companies shared some of the gains with workers, consumers, and shareholders. Less cash to the government would mean more cash to spread around to everyone else. Indeed, the research by Hanlon, Hoopes, and Slemrod indicates that immediately after the law passed, hundreds of companies announced plans to raise wages and salaries, issue bonuses, or hire new workers.

Kennedy, Dobridge, Landefeld, and Mortenson looked at a random sample of federal tax records for both workers and companies and find that while the TCJA did increase average C-corp payrolls (in total dollars) by 1.2 percent, these higher payrolls were largely driven by the top employees. The gains ended up adding 5 percent to the compensation of upper-income employees and executives. Nine out of 10 members of this group were men, with an average age of 53 and annual earnings of over $1 million.

Breaking down the gains in dollar terms, $55 billion went to business owners, $11 billion to executives, and $32 billion to employees in the top 10 percent of their companies' wage distributions, according to the researchers. The TCJA provided the remaining 90 percent of the labor force with no pay hikes at all, they conclude. "Overall, the results imply that corporate tax cuts improve aggregate efficiency but exacerbate inequality," they write.

The TCJA provided 90 percent of the labor force with no pay hikes at all, researchers conclude.

Party like it's 1997

Even if the predictions made by the TCJA's backers are wrong, lawmakers aren't eager to reverse the cuts. Perhaps to make the 10-year cost of the law more palatable, the TCJA's authors established that a number of provisions, including some relating to individual tax rates, would expire in 2025. But there would be political ramifications to raising taxes on more than half of Americans, including middle- and lower-income ones.

Yet making all the individual provisions permanent for every income group would, according to the Congressional Budget Office, have a price tag of $3.5 trillion over the next decade. So what to do?

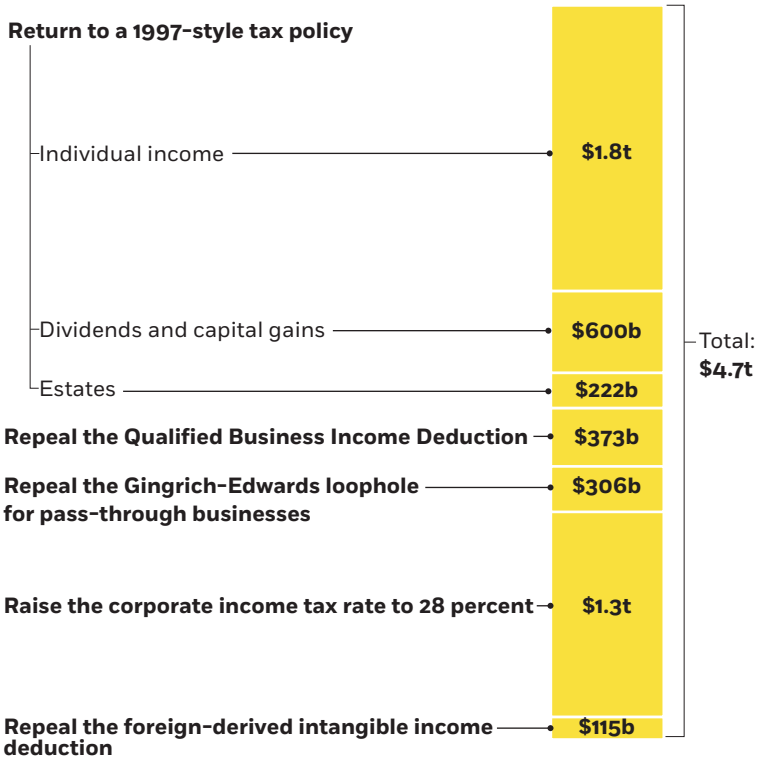
Zidar and Zwick recommend that lawmakers take a dispassionate look at each provision and decide which to keep. They also argue that as lawmakers formulate what the tax code will look like beyond next year, they should return to a 1997-style tax regime—or, as the researchers put it in their Aspen paper, "party like it's 1997."

In some ways, the current climate is reminiscent of that year, they point out. The late 1990s was another time of divided government: in January 1997, Bill Clinton began his second term as president while House speaker Newt Gingrich (Republican of Georgia) and Senate majority leader Trent Lott (Republican of Mississippi) oversaw Congress.

A few ways to raise revenue

The researchers suggest reforms that include reverting some tax rates to 1997 levels.

Suggested tax reforms* and 10-year revenue estimates



*Revenue estimates for returning to a 1997-style tax policy and repealing the Qualified Business Income Deduction come from the Penn Wharton Budget Model. The rest are from the US Dept. of Treasury.

Zidar and Zwick, 2023

But the tax rates have markedly changed since then, even as the political battles still wage. The tax code at that time had higher rates on estates, on what top earners paid on dividends and capital gains, and on individual income. Sidestepping the politics of such proposals, the researchers looked instead at what academic literature suggests about each component.

Going back to 1997's rates for individual incomes would raise $1.8 trillion over a decade, per the Penn Wharton Budget Model, which provides nonpartisan estimates and analysis of US legislation.

Given the many levers the affluent can use to minimize taxes, as well as the interplay between business and individual tax rates, the only way to meet the US's pressing financial needs may be to hike the top marginal tax rates applied to income from business activity, the researchers argue. Returning to the January 1997 regime would result in high-earning married couples paying 36 cents instead of 24 cents on income above their first $300,000. For those making $500,000, the rate would go up to 39.6 percent from 35 percent. At the opposite end of the income spectrum, the researchers propose relief akin to the Making Work Pay Tax Credit that was part of the American Recovery and Reinvestment Act of 2009.

"There are a host of ways to avoid ordinary income taxes by deferring income into a form classifiable as capital gains, such as carried interest, qualified small-business stock, and incentive stock options," write Zidar and Zwick. "In our view, these carve-outs generally allow individuals to delay compensation and enjoy a lower tax rate on what is often labor income in its underlying nature. Since much of this activity is labor income, it should not be tax-advantaged relative to that of wage earners."

Estate taxes are another area where Zidar and Zwick favor 1997 rates. Currently, a rate of 40 percent kicks in on inheritances in excess of $13 million. But in 1997, the estate tax rate was 55 percent on inheritances in excess of a little over $1 million. This, combined with a repeal of the cost-basis "step up" that heirs currently receive—and that relieves them of paying taxes on previous gains—would raise $222 billion over 10 years, according to the Penn Wharton Budget Model.

The PWBM also calculates that changing tax rates on dividends and capital gains back to earlier levels would raise $600 billion over 10 years. Yale's Natasha Sarin and Harvard's Lawrence H. Summers, with Zidar and Zwick, have argued that raising capital gains rates to match ordinary-income rates could add another few hundred billion dollars to

Zidar and Zwick are also nostalgic for 1997 when it comes to funding for the Internal Revenue Service.

that tally. (For more, read "Could the US raise $1 trillion by hiking capital gains rates?" in the Summer 2021 issue and online at **chicagobooth.edu/review**.)

Zidar and Zwick acknowledge the argument that tax increases could hurt economic growth, but they note that growth was swift in the late 1990s. They also point to research by Yagan, who studied the effects of a 2003 tax cut on dividends by comparing similar enterprises that were structured as C corporations, and thus qualified for the dividend tax cut, with S corporations, which did not. (For example, Home Depot is a C corp while Menards is an S corp.) Yagan's research finds that the dividend tax cuts had no effect on a company's investment. So reverting to a 1997-style system of taxing dividends at the top individual tax rate would have limited effects on competitiveness and economic growth, Zidar and Zwick conclude.

As for capital gains, most Americans pay a higher tax rate on their salaries and wages than on capital gains, a policy that benefits people who own business assets, who tend to be relatively wealthy.

According to Yagan and his colleagues Emanuel Saez and Gabriel Zucman at UC Berkeley, 42 percent of unrealized capital gains take the form of private business gains. What's more, the higher up the wealth spectrum, the bigger the share of personal assets that are held as private business equity. Among centimillionaires (worth at least $100 million), two-thirds of unrealized capital gains are in the form of private business equity.

How much would raising capital gains rates damage economic activity? Tax fights in the 1990s featured dueling research findings on this, recount Princeton PhD student Ole Agersnap and Zidar. "This issue has reemerged in every presidential administration since 1990 and plays a key role in ongoing tax reform plans. For instance, this elasticity is the central parameter governing the revenue scores of President Joe Biden's plan to increase capital gains rates as well as President Trump's proposal reducing capital gains taxes," they write. For their part, Agersnap and Zidar looked at state-level data to estimate how state capital gains tax changes affected where wealthy Americans lived and how they realized their capital gains. They then built a framework to estimate how the patterns would play out nationally and find that the economic response of capital gains realizations to changes in capital gains is likely modest, on the order of between -0.3 and -0.5 over 10 years. (Thus, for every 1 percent rise in the rate, realizations fall by 0.3 to 0.5 percent.) For context, the researchers cite other research estimates that range from -3.8 to -0.22.

Zidar and Zwick are also nostalgic for 1997 when it comes to funding for the Internal Revenue Service. The IRS budget as a share of GDP was almost 0.09 percent in 2002 but closer to half that in 2020—with a corresponding decline in audit rates.

Even deeper in the weeds of tax policy, Zidar and Zwick point to additional changes that could yield big benefits for the Treasury. These include repealing the so-called Gingrich-Edwards loophole, which allows taxpayers to characterize income from consulting and speaking fees as business profits rather than wages. Per the Treasury Department, this would raise $306 billion over 10 years.

Another option: allow the TCJA's Qualified Business Income Deduction to expire. This deduction has lowered tax rates on many pass-throughs, and scrapping it would raise $373 billion over 10 years without a big effect on investment or growth, the PWBM estimates. Zidar and Zwick offer other recommendations, as well, all of which they say add up to $4.7 trillion.

The politics of taxes

Predictably, not everyone is on board. For one thing, Zidar and Zwick aim to increase tax progressivity. Democrats tend to focus on sharing the economic pie equally by implementing "progressive" policies, while Republicans are more focused on expanding the pie through policies they believe will spark economic activity. One thing both sides agree on is that their policies are the best way to benefit rank-and-file workers and their families.

John Cochrane, a senior fellow at the Hoover Institution, says he'd take an entirely different approach. It would include eliminating corporate taxes altogether, and replacing levies on incomes and estates with a broad-based consumption tax. (Read more in "It's time the US abolished the income tax," Spring 2024 and online.)

The politics of this discussion are even more complicated than the math.

The US regime is already one of the most progressive in the world, says Hoopes, who is research director of the UNC Tax Center. He notes that while the US was ratcheting down income tax rates for high earners in recent decades, Americans at the other end of the spectrum were benefiting from provisions including the Earned Income Tax Credit and the Child Tax Credit. The result is that about half of all earners pay no federal income taxes at all, although they do pay payroll taxes. What's more, says Hoopes, no matter how Washington hikes taxes on top earners, it can't tame its deficit problems without coming up with other revenue sources or spending cuts.

The politics of this discussion are even more complicated than the math. President Barack Obama proposed some of the same corporate tax changes that were ultimately passed under Trump, and while some Democrats supported Obama's proposals, they disparaged them when advanced by his successor, Hoopes recalls. However, he also says that with divided government, Congress may be forced to forge practical solutions that have enough support to pass and that provide the longevity the private sector needs to operate efficiently.

Tax policy is sure to remain politically contentious. This polarization is epitomized by a pair of opinion columns from this spring. One in the left-leaning *New York Times* blames low taxes for expanding America's wealth gap and the ranks of its billionaire class. Another in the conservative *Wall Street Journal* declares that "The U.S. Already Soaks the Rich," citing a study that indicates the top 1 percent of earners already pay close to half the nation's income taxes.

The 2025 deadline written into the TCJA makes another tax fight practically unavoidable, and the outcome will be crucial to the federal government's solvency. To raise revenue, it isn't necessary for lawmakers to tear up the tax code and start over, Zidar and Zwick argue. A return to 1997 could be the answer.–CBR

Go to **chicagobooth.edu/review** to see citations for research mentioned in this article.

COMPANIES SHOULD FOCUS ON PROFITS.

SHOULDN'T THEY?

Booth researchers are exploring what should drive corporate decisions in the 21st century.

Expand your thinking at
chicagobooth.edu/company-profits

ANSWERING QUESTIONS AND QUESTIONING ANSWERS.

CHICAGO BOOTH
THE UNIVERSITY OF CHICAGO BOOTH SCHOOL OF BUSINESS

FOOTNOTES

Look beyond your own experience **Page 58**

Healthcare and the moral hazard problem **Page 60**

JANE L. RISEN

An argument for less debate

For better understanding and decision-making, try dialogue instead

One of the University of Chicago's most closely held and well-known values is its devotion to rigorous inquiry and free and lively debate. On the university's website, we say that "an education with free and open debate empowers students to grapple with challenging ideas." Indeed, I can't think of many other institutions as committed to having a marketplace of ideas where those ideas can compete.

ILLUSTRATION BY MATT CHASE

But in the spirit of rigorous inquiry, I'd like to question the notion that debate is always the best way to share our ideas or discover truth. Instead, I'd like to make a pitch for dialogue.

Like any good debater, I'll start by offering some definitions. When I say *debate*, I mean any time that we disagree with someone and engage that person with the goal of persuading her that our point of view is correct. We can debate international affairs or the proper way to load a dishwasher. If the goal is to convince the other person, we'll think of that as debating.

In contrast, in *dialogue*, the goal is understanding one another. My goal is to understand you and to have you understand me. Likewise, your goal is understanding and being understood.

Debates tend to feel competitive, with people assuming that if one person is right, the other must be wrong. In contrast, dialogue often feels collaborative. In debate, I'm focused on poking holes in your argument. In dialogue, though, I'm more likely to ask questions to make sure I really understand what you're saying.

We can think of a constellation of goals, assumptions, and behaviors as being more debate-like or more dialogue-like. In the past couple of years, my students, collaborators, and I have been exploring the consequences of engaging in either debate or dialogue, as well as the factors that lead people to spontaneously engage with a more debate-like or dialogue-like approach.

The origin of my interest in this topic—in fact, the root of my interest in social psychology more broadly—is an opportunity I had as a teenager to participate in a conflict transformation program called Seeds of Peace. The program brings together teenagers from regions of conflict around the world to spend three weeks together at summer camp.

In the early years, including the time I spent there, Seeds of Peace focused on teens from the Middle East, and I was fortunate that they also invited a handful of Americans. Watching people who had been taught to be enemies their whole lives get to know one another as individuals was truly a remarkable and life-changing experience.

In many ways, the program operates like any summer camp. Kids play sports, live in cabins, and sit around the campfire. But in other ways, it's unusual. Specifically, campers spend almost two hours a day in dialogue sessions with a designated group of fellow campers. In dialogue, they tackle the most painful and divisive issues defining their conflict, share their personal experiences, reflect on competing narratives, and challenge each other's prejudices. No subject is off-limits. These sessions are intense and emotionally exhausting for campers.

Maybe more importantly, though, let me tell you what dialogue is not. As described on the program's website, "The purpose of dialogue is not to come to consensus or agreement, but to more deeply understand the differences that each individual brings, to listen and to be heard in all of our complexities, and to learn something new about oneself and others."

In addition to sparking my passion for psychology, Seeds of Peace became an eventual research partner. In one recent paper, my coauthors and I explored the relationships that form at this camp. In general, we find that campers were more likely to form connections with those who shared their nationality. In other words, those who were part of their in-group rather than those in the out-group. This is consistent with lots of evidence showing the power of similarity, or what we call homophily: liking people who are like you.

But remarkably, when we specifically examined the relationships that formed among campers who shared a dialogue group, the pattern fully reversed. Campers were more likely to form relationships with out-group members in their dialogue group than in-group members.

Let that result sink in. This means that a Jewish Israeli camper was more likely to become close with a Palestinian camper in his dialogue group than with another Jewish Israeli in that group. And likewise, a Palestinian was more likely to become close to a Jewish Israeli in

For a task with an objective right answer, we find that pairs were more likely to get the correct answer when engaging in dialogue.

To foster dialogue, we should embrace humility, prioritize learning, and intentionally look for opportunities to understand and appreciate those who are different.

the group. I don't know that there is another finding that I could share that would better highlight the remarkable power of dialogue.

Now, imagine instead that the Israeli and Palestinian campers engaged in two hours of debate every day for three weeks. I feel quite sure that we would not see those same relationships form.

We couldn't test that hypothesis at camp, of course, but we could in the lab. There, we randomly assigned pairs of participants who disagreed about an issue to have a conversation in which they were instructed to focus on either demonstrating why they were right or sharing and learning about each other's perspectives.

For a task with an objective right answer—choosing the best candidate on paper to hire—we find that pairs were more likely to get the correct answer when engaging in dialogue. When the interaction was an online chat, we find that participants were more satisfied and felt more included in dialogue than in debate. Those in the dialogue condition also had a more accurate understanding of their partner's perspective.

What about persuasion? Did those in the debate condition manage to convince each other that they were right? No. If anything, participants seemed to change their opinion more after engaging in dialogue. Thus, these initial findings suggest that dialogue can improve objective decision-making, subjective experience, and understanding.

Given the apparent benefits from engaging in dialogue, we also wanted to understand what leads people to approach disagreement in a more debate-like or dialogue-like way. We find that when pairs more strongly disagreed about an issue and when individuals felt more certain of their own opinion, they were more likely to engage in debate and less likely to engage in dialogue. Participants were more likely to engage in dialogue, however, when they perceived that they shared goals and values with the other person.

With these findings in hand, I'd like to highlight three things we can all do to encourage more productive disagreement:

First, be humble. We need humility to recognize that we haven't figured everything out yet. When we have humility rather than certainty, we can make space for other good ideas and perspectives.

Second, adopt and pursue learning goals in both our personal and professional lives. When we pursue learning goals, we can put winning aside—at least sometimes—and instead focus on growing and improving. With a learning and growth mindset, we are better prepared to learn from successes as well as failures. We are free to make mistakes and admit to them because we recognize that we are a work in progress just like everyone else is.

Finally, be intentional. There's something powerful that comes from recognizing our default patterns. Once we know that people tend to connect more easily with those who are similar and that people are more likely to start a debate with those who seem different, we can intentionally disrupt our default patterns. If we're intentional about learning, and especially if we're intentional about learning from people who have different perspectives, we'll be in the best position to capitalize on all of the best ideas and continue to grow.

To foster dialogue, then, we should embrace humility, prioritize learning, and intentionally look for opportunities to understand and appreciate those who are different. If we can all engage in a little more dialogue and a little less debate, I believe the marketplace of ideas will grow more vibrant—and more people will want to spend more time shopping there.–CBR

Jane L. Risen is the H. G. B. Alexander Professor of Behavioral Science and a John E. Jeuck Faculty Fellow at Chicago Booth. This is an edited transcript of the speech she gave this past spring at Booth's 2024 Graduation Ceremony for the Evening, Weekend, and Executive MBA Programs.

RAGHURAM G. RAJAN

ROHIT LAMBA

Democracy and innovation could set India on a different development path

The emphasis should be on new companies, ideas, and products that allow the country to own the high end of the value chain

India can adopt a new path to development, one that no developing country has taken before, wherein its firms come up with world-beating ideas and products and deliver them globally.

No large developing country has skipped the middle step in the typical development route, which entails first shifting workers from agriculture to manufacturing before shifting them again to services. India has partly jumped from agriculture straight to services, but it must now reinvent itself once again so as to accelerate growth and provide jobs to the teeming millions joining its labor force every year.

Instead of making generic pharmaceuticals, which it has been good at, India should turn to finding new cures for the diseases that plague its people and sell those new medicines to the world. Instead of buying expensive 5G technology from a vendor in an industrial country, India should create a cheaper version domestically and sell it to the emerging world, assuring buyers that India will create no backdoors through which it can snoop on them. It is important to recognize that India has the foundations on which it can build to fulfill these aspirations. But it is not there yet.

For instance, India has only a few top-quality research institutions, such as some of the Indian Institutes of Technology, the Tata Institute of Fundamental Research, and the Indian Institute of Science. To move from incremental innovation to pathbreaking innovation, India needs to raise many more of its universities to global standards, and encourage and fund innovative research as well as business-academia collaborations.

India should also create the conditions for more manufacturing at home, but the export-led, low-skilled variety—such as the assembly of electronics components or the stitching of garments—has become highly competitive and no longer offers an easy path to becoming a middle-income nation. Instead of trying to capture the bottom of the value-added chain and climbing up from there, as the East Asian countries did, India could aspire to own the high end of the value chain directly. In some cases, the low-skilled segments would then migrate naturally to India. While high-skilled services can

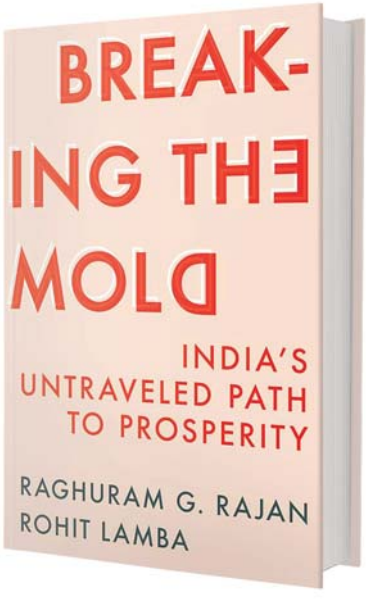
also expand to provide the foreign exchange and jobs India needs, the emphasis should be on new firms, ideas, and products, whether in manufacturing or services, that can allow India to leapfrog.

The right climate for innovation

One critical support to India's development path will be its democracy. Citizens benefit intrinsically from democracy—the dignity that comes from being able to vote and possessing the right to express your opinion through your ballot, having freedom of thought and expression more generally, being treated fairly, enjoying the rule of law, and so on. India's citizens just exercised their universal adult franchise in the recently conducted national elections, where about 650 million people voted. But there is also an instrumental reason India should strengthen its democracy.

In the early stages of development, the focus is, as we have seen, on catch-up growth. The ideas and know-how needed for development are already out there, discovered by some other country and its businesses; they simply have to be imitated or licensed.

The development path we suggest will depend far more on Indians having innovative ideas and being creative, pushing the intellectual frontier. Growth at the frontier requires debate and argumentation, which an authoritarian government rarely tolerates. It is not that authoritarian countries cannot innovate to some degree—the Soviet Union had a flourishing military-industrial complex. But authoritarian governments want to direct research and innovation, which ensures they are limited by the imagination of those in charge. And when those in charge are apparatchiks, research and innovation will be limited indeed, especially if the apparatchiks interfere constantly because they worry these directions may not be consistent with the views of the supreme leader.



By contrast, innovation in a democracy does not have to respect the existing power structure and its beliefs, and thus can be really pathbreaking. Chip technology in the Soviet Union always lagged that of the United States because the Soviets simply could not innovate in that area. They fell into a pattern of trying to steal the intellectual property when it became widely available, which constantly put them behind.

Similarly, China's political system may have been ideal for catch-up, infrastructure-led growth. Arguably, it is much less so as China approaches the technology frontier.

For instance, China has imposed requirements that artificial intelligence essentially respect the primacy of the Communist Party. This could limit the extent to which firms can explore what AI can do, for fear they might inadvertently cross regulatory boundaries. In the *Financial Times,* Yu Jie, a senior research fellow at Chatham House, argued that China's ability to make major scientific breakthroughs will depend "on whether researchers have the space to think critically and creatively."

Careful studies by economic historians have addressed the relation between creativity and economic and political freedom over the long run. One study—by University of Mannheim postdoctoral researcher Alexander Donges, University of Texas at Dallas's Jean-Marie Meier, and Rui C. Silva at the Nova School of Business—looks at the consequences of the French occupation of German counties after the French Revolution in 1789. Among the key reforms the occupying French implemented was abolishing local guilds (an early form of business cronyism). The French also brought

in their civil law, under which the judiciary was made independent from the local administration and all citizens were treated equally before the courts. Counties that were occupied for the longest period, and thus saw these French reforms take greatest hold, had more than twice as many patents per capita almost a century later, in 1900, than counties that were not occupied, according to the study.

A related study by University of Essex's Michel Serafinelli and Bocconi University's Guido Tabellini looks at which cities in Europe saw a rise in notably creative people, whether by birth or by immigration, between the 11th and 19th centuries. The researchers find that independent cities that assured their citizens political freedoms were most likely to see a rise in the presence of such people, because the city environment fostered creativity and also because such cities attracted creative souls.

This suggests the arguments made by some that all will go well in India if only it has an iron-fisted leader are not empirically well grounded. Today, such an authoritarian leadership can only build more roads and monuments, running roughshod over people's rights, but cannot allow the environment of free thinking and speech that India needs for innovative ideas and products—for such a liberal environment will open the door to criticism of all authority. By trying to control debate, any authoritarian government will make it hard for India's researchers to be innovative and for its institutions to attract free thinkers from the diaspora or retain its youth, who are unhappy with the status quo. Policing of thought is certainly not what India needs today.

Is manufacturing needed for security?

What about manufacturing's role in national security? If India does not have a strong manufacturing base, will its national security be impaired? Of course, it is important for a large country to have its own domestic defense industry so that it is not subject to undue outside pressure in times of increasing conflict. But does it have to manufacture every part that goes into every weapon?

What if the unimaginable happens and the democratic world turns against India? If so, simply having factories making older chips will not be enough.

Take, for example, the global race to get into the manufacture of advanced logic chips. Both the US and Europe are trying to bring more high-end chip fabrication to their shores, while China is trying to upgrade its existing facilities as the US bans the sale of high-end chips to China. Should India also fabricate chips?

Start first with the obvious point, that if short-term disruptions in chip availability are the concern, as during the pandemic, the solutions are simpler. India could incentivize firms to have larger inventories of critical chips, and even possibly create a small national reserve. It could source chips from multiple countries and companies. Indian firms could build flexibility around production processes so that products can be redesigned to replace the chips in short supply with the chips that are available. All of this is much cheaper than manufacturing chips domestically, given that even a plant making chips that are a few generations behind the current technological frontier will cost tens of billions of dollars in subsidies.

If the longer-term concern is that India might face sanctions from potential enemies, the solution is to have a wider and more diversified set of friends. It is hard to imagine that democratic India will take a course of action that will make the euro area, the US, South Korea, Japan, and Taiwan all want to sanction it. But what if the unimaginable happens and the democratic world turns against India?

If so, simply having factories making older chips will not be enough. India will need to make state-of-the-art chips (that is, the kind that go into mobile phones and AI machine-learning processors); it will need to make the machines that make the chips (firms that make those machines, such as the Dutch company ASML, will also apply sanctions); and it will have to make every part of the chip supply chain, starting from the silicon wafer—all of which require specialized processes and chemicals that India does not have.

Put differently, unless India brings the entire manufacturing process for chips into India, there will always be choke points that run through other countries. Total self-sufficiency is nearly impossible, even if India

is prepared to invest hundreds of billions of dollars.

In short, India cannot obtain security with a toehold in chip manufacturing. Chip manufacturing, unless more carefully thought out, could be like the prestige-project white elephants that India has had plenty of in the past. Should India spend tens of billions in subsidizing chip manufacturing when the world has periodic gluts of chips, or should it devote those tens of billions to opening tens of thousands of high-quality primary schools, thousands of high-quality high schools, and hundreds of top-notch universities? Is India better off dominating chip design with the tens of thousands of additional engineers and scientists it will produce, and starting firms like US-based Nvidia, Qualcomm, or Broadcom, none of which fabricate their chips? Or does India want to imitate China, especially when it has much better relations with the chip-manufacturing world? Once again, rather than following others blindly, India needs to look at its own advantages.

Some argue that India needs chip fabrication so that it can build strength in chip design or other parts of the supply chain. There is no evidence that this is the case—witness Nvidia or ASML. That other countries or regions are jumping to subsidize chip fabrication is good for India; it will increase its choice even if it does not produce, especially when the periodic chip glut emerges.

That is not to say India should never enter chip fabrication. As the current frenzy of subsidies dies down, investment in the industry will, eventually, be worthwhile. India's trained engineers and designers will have the human capital to participate in the innovation that is so crucial in this industry. India should not hesitate at that point. Nor should it hesitate if anyone wants to invest in India without massive subsidies. But it does not seem wise at this moment to enter this ruinous subsidy game—India is better off investing in its human capital.

The untraveled path

While ideas and creativity should be India's main vehicle for growth, there is a reason services and manufacturing-related services exports may be easier for India to expand in than manufacturing exports. China's dominance in the area, and the consequent loss of middle-income factory jobs in industrial countries, has made the West wary of manufacturing imports. Protectionism in manufacturing is rife, as is competition for the remaining shrinking pie. There is little room for another China-sized exporter of manufactured goods. This is not to say that India should not manufacture goods, only that it should be aware of the new limits to China-style growth. It should think especially hard if incentivizing such manufacturing requires taxpayer subsidies and protectionist tariffs paid by domestic customers.

An Indian consultant still costs a fraction of what a US consultant does, with pretty much the same capabilities, albeit a somewhat different base of experience.

Services, however, are still relatively virgin territory. For instance, an Indian consultant still costs a fraction of what a US consultant does, with pretty much the same capabilities, albeit a somewhat different base of experience. This is why many global service firms are looking to India even without any subsidies.

As the world grows richer and older, it will expand its use of services, but climate change suggests additional urgency. The world has to slow the growth in its consumption of goods. This offers one more reason for India to be mildly biased toward services and manufacturing-related services, even while emphasizing that its focus should be on ideas and creativity, wherever these apply. In sum, the notion of premature deindustrialization, as written about by Harvard's Dani Rodrik, need not be a bug but a feature of India's growth path.—CBR

Raghuram G. Rajan is the Katherine Dusak Miller Distinguished Service Professor of Finance at Chicago Booth. Rohit Lamba is an assistant professor of economics at Cornell.

Go to **chicagobooth.edu/review** to see citations for research mentioned in this article.

ILLUSTRATION BY FRANCESCO CICCOLELLA

DOUGLAS W. DIAMOND

Look beyond your own experience

Let the events of your life enrich, not bias, your thinking

Every graduating class has its own distinctive suite of personal experiences that helps to shape its perspective. For the business school students who graduated this June, for example, COVID was a big part of those experiences; most of them were already employed in 2020, when the virus upended the world, and therefore had the opportunity to see how their companies responded to the pandemic and how that response played out. Then graduate school gave them the time and environment to reflect on what they'd seen and how it affects their views of the world. Their teachers, if they did their jobs well, forced them to develop, defend, and discuss many ideas related to business, including ideas based on that recent personal experience.

What you've experienced in the past has a profound influence on your expectations for the future. That notion has been well established in research by my Chicago Booth colleague Stefan Nagel. Nagel studies finance, but this line of research is based on a pretty big literature in psychology. It's about the effects of your personal experience—your day-to-day life, the things that happen while you're alive and that you didn't just read about—on your beliefs, expectations, and predictions of the future.

Nagel studies this in two dimensions. One is an individual's personal experience with inflation. He and his coauthors looked at predictions about the level and persistence of inflation. And presumably it's fairly easy to learn about inflation that happened when you were two years old, or five, or before you were born, but that has a different impact on your beliefs about inflation than does inflation that you actually experience. The researchers demonstrate that there is a huge difference between individuals who are otherwise the same but who experienced different inflation during their lifetime.

Even experts are susceptible to this. In separate but related research, Nagel demonstrates that central-bank governors—who are somewhat educated about past inflation, or at least we would hope so—are influenced in their monetary-policy voting decisions by their personal inflation experiences.

Nagel also looked at how the performance of the stock market during your lifetime predicts how big a fraction of your investment portfolio you'll put into equity. And sure enough, if stocks do well, you'll likely put more into equity. If stocks did poorly in the past—in *your* past—you'll likely put less in.

In short, the evidence is strong that personal experience is a big deal for the expectations you form and, in turn, the decisions you make.

One role of education is to help overcome the bias of personal experience. When students are exposed to other people's experiences and decisions, and to ideas from faculty in wide-ranging disciplines, it can help them to counter the tremendously overweighted role of their own personal histories.

A former dean at Booth who's now at Yale, Edward A. Snyder, told incoming MBA students: "You should complain vociferously if our research-oriented faculty are not doing research. And even more than that, not bringing their research into the classroom." I think what he meant was that there are benefits of bringing together people with different perspectives and different views who have thought carefully about the same issue. It's good advice.

Seek out and incorporate the perspectives of others.

Thinking about this as a teacher and researcher, I can say that discussion of these ideas in the classroom doesn't just benefit students. Two of my own papers that I've learned the most from writing came out of teaching MBA students in my financial markets and institutions course and being unable to answer the question, "Why did this happen?"

I hesitate to generalize, but the most successful Booth alumni I know are those who implemented ideas that were very, very far from the norm in their industry. Being able to draw on different views, different beliefs, and different perspectives is important for innovation. When your thinking isn't confined to the events of your own experience, you've got a major competitive advantage in an environment where you've got to get out first ahead of rivals.

Of course, personal experience is not in itself a bad thing. In fact, you may need some personal experience to understand and appreciate what you've learned in the past. When I hear from my former MBA students, it's usually during financial crises, after they've experienced the issues that I cover in my class. They often provide interesting examples of how these things played out in their business and how they maybe were a little more prepared than others to figure out what to do.

Your experiences matter more than you may realize. The key is to use them to amplify and build on your knowledge, rather than being biased by them. Seek out and incorporate the perspectives of others, and you can benefit from the value of your experiences without giving them undue weight.–CBR

Douglas W. Diamond is the Merton H. Miller Distinguished Service Professor of Finance at Chicago Booth and a 2022 recipient of the Nobel Prize in Economic Sciences. This is an edited transcript of the speech he gave this past spring at Booth's 2024 Graduation Ceremony for the Full-Time MBA and Stevens Doctoral Programs.

TAL
GROSS

MATTHEW J.
NOTOWIDIGDO

Healthcare and the moral hazard problem

The demand curve isn't simple when lives are on the line

A 70-year-old woman goes to the pharmacy to pick up medication for her arthritis. How much should that cost her? Maybe $5 for the prescription? Or $20? Or should it be free?

There's surely a lot going on. The woman might be grappling with poverty and discrimination and a plethora of challenges beyond arthritis. A pharmaceutical company may have priced the medication very aggressively. The woman might not have access to good medical advice. For now, though, we want to focus on one, narrow question: How much should she have to pay for her medical care?

Health economists have grappled with that question for as long as there have been health economists. The answer is not simple, and the debate continues to this day. The short answer: it depends. The long answer requires a tour of research on the issue, research that goes back nearly half a century.

The root of the problem

High prices are awful. No one likes it when things are expensive. What kind of *monster* would like high prices?

An economist.

High prices do something important: they force people to agonize over whether or not they really want to make the purchase. It can be problematic for people to consume goods without having to grapple with their price.

That's why, whenever the government wants to reduce the consumption of something, there's a simple solution: just raise the price! If there's too much traffic in a city center, raise the price on driving downtown through higher toll prices or "congestion charges." If there's too much pollution, raise the price on pollution through a carbon tax. If too many people smoke, raise the price of cigarettes through tobacco taxes.

Why turn to higher prices? Because higher prices lead people to align their personal decisions with the full cost of production, whether those prices reflect the costs we typically think about or harder-to-measure costs such as congestion and pollution. They lead people to consume less.

What does this have to do with healthcare? Health insurance fundamentally breaks the relationship between individual decisions and the costs of production. It breaks that relationship because, by definition, generous health insurance shields consumers from the high price of healthcare. A problem with generous health insurance is that it makes healthcare too cheap. And when healthcare is too cheap, people buy too much of it.

Health insurance, in other words, can eliminate the benefit of high prices, the way that they force people to grapple with the costs of production. And that can lead to waste. If a 70-year-old woman doesn't have to pay anything for her arthritis medication, she might continue with the medication even if it's not working. That medication still costs the healthcare system money, money that could be better spent on more effective forms of healthcare.

There are plenty of contexts in medicine in which there's a cheap option and an expensive option. Sometimes, the cheap option is just as good as the

expensive one. For instance, an upper respiratory infection can be treated in an emergency room (expensive) or in an ordinary doctor's office (cheap). Some conditions can be treated with generic drugs (cheap) or branded drugs (expensive). If the consumer pays the same price for either option, why not choose the expensive option? And if all consumers face the same incentives and behave in the same way, healthcare spending, overall, might rise in ways that don't actually improve health.

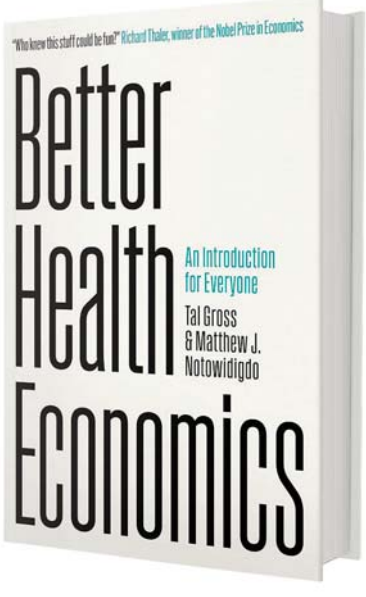
There's a technical term for this issue: it's called a "moral hazard problem." In general, moral hazard problems are situations in which there are two parties in a transaction and one party cannot control the actions of another. In the case of health insurance, there's the insurer and the consumer, and the insurer bears the costs of the consumer's healthcare decisions. The consumer can choose the cheap option or the expensive option, and the insurer pays either way.

Now, to be clear, that's not to say that generous health insurance is a bad thing. Generous health insurance also protects consumers from risk. The issue here is that there's a trade-off. On the one hand, we want health insurance to be generous so that people are protected from risk. On the other hand, we *don't* want health insurance to be generous, because of moral hazard. The generosity of health insurance—how much the consumer has to pay for healthcare—has to balance those two forces.

But all of that is *theory*, words on a page that describe how some people *think* the world works. Theory needs to be tested. Next, let's turn to real-world evidence on moral hazard in health insurance.

What do copayments do?

Health economists first examined moral hazard in health insurance with a field experiment. In the late 1970s, a team of health economists sat down at the RAND Corporation, a think tank in Santa Monica, California. Joseph P. Newhouse, now at Harvard, was a young economist just starting out at RAND. He asked his colleagues a simple question: How do people respond to the price of healthcare?



If ordinary people have to pay more for healthcare, do they consume less of it? Newhouse found that lots of economists had opinions on that question, but no one had any good evidence.

Some argued that healthcare was different than other goods, that healthcare is always a matter of "your money or your life." Therefore, they argued, people would pay whatever price they faced for healthcare—the price didn't matter, because healthcare was so important. And, by extension, those people were not concerned with moral hazard: since healthcare is different from other goods, it doesn't matter, they argued, that health insurance makes healthcare cheap.

Economists knew that prices matter for ordinary goods: coffee, wheat, motorcycles. If the price of coffee goes up, people buy less coffee. That is, as we say in Econ 101, a matter of the demand curve. Is the same true for healthcare? No one knew.

The RAND researchers, led by Newhouse, decided to run an experiment. They took 2,750 American families and randomized them into two groups. The study included both urban and rural households and spanned a broad range of income levels. One group of families was put on a free-care plan: for several years, all of the healthcare they needed would be free. Every doctor visit, every dentist visit, every medication: they would pay nothing.

Other families were put on a high-deductible healthcare plan. They would be responsible for all of their healthcare costs up to a thousand dollars. After $1,000, their health-insurance plan would kick in and cover everything. But until then, they had to foot the bill on their own.

Remember: randomization makes experiments valid—it means that both groups began the experiment with the same health, on average. They had the same average income, level of education, number of children, number of televisions, and, most importantly, the same average health. As a result, any differences in outcomes in the years following the RAND experiment can be interpreted as the impact of the health-insurance plans themselves.

The study became known as the RAND Health Insurance Experiment and it lasted from 1976 until 1982. For health economists, the experiment amounts to a combination of NASA launching a space shuttle, Bill Gates starting Microsoft, and Ayatollah Khomeini returning to Iran. It happened in the late 1970s, and it's a big deal to us. The RAND experiment is one of the most expensive experiments ever performed by social scientists.

For years, the families participating in the experiment led their ordinary lives; the only thing out of the ordinary was that a team of researchers at RAND was handling their health insurance. Then, after years of being on either the free-care plan or the high-deductible plan, the participants were given a final physical exam, and the experiment was over.

Three long-term takeaways

Newhouse and his colleagues spent years poring over the data, trying to understand how having to pay for healthcare affected families. The researchers studied the results from every possible angle, slicing and dicing the data every which way. The results of the experiment filled hundreds of academic papers and also a 516-page book. But, decades later, the most relevant discoveries from the experiment boil down to three main conclusions:

High prices really do cut how much care people choose to consume. First, the researchers compared the amount of healthcare consumed by families that were put on the free-care plan with that of families randomly assigned to the high-deductible plan. Those put on the free-care plan consumed an average of almost $2,000 (in 2021 dollars) as compared with about $1,600 in healthcare for those on the high-deductible plan. That's a roughly 20 percent difference—a large difference.

In other words, incentives matter, even for healthcare. People who face a higher price for healthcare consume less healthcare. In the language of Econ 101, demand curves slope down, even for healthcare. Yes, healthcare is important, and people treat it as important, but, at the end of the day, the price still matters.

Deductibles lead families to cut back on all healthcare, regardless of whether it's effective or ineffective. The second conclusion of the experiment arose as the researchers tried to figure out *which* healthcare the families on the deductible cut out. The researchers assembled a panel of physicians and gave them all of the medical charts associated with the experiment. They asked the physicians to categorize all of the healthcare as "highly effective" care or "rarely effective" care. Going to the ER for a runny nose: that's rarely effective care. Going to the ER for a heart attack: that's highly effective care.

The panel of physicians worked through the stack of charts, methodically categorizing all visits as highly effective or rarely effective. Then they studied how the high-deductible plan affected those two categories.

Families facing a deductible cut back on highly effective care by about 30 percent relative to those on the free-care plan. And then the researchers found roughly the same effect for rarely effective care: a roughly 30 percent drop in utilization. The lack of contrast between those two findings is the second conclusion of the experiment.

And that finding alone is kind of disappointing. Health policy would be much simpler if people behaved like shrewd medical experts whenever they faced a deductible. Unfortunately, that's not how it works. Patients are not physicians themselves—they don't know what is effective and what is ineffective. When faced with a high price, they just cut back on all of it, both care that really matters and also care that is probably wasteful.

Health policy would be much simpler if people behaved like shrewd medical experts whenever they faced a deductible. Unfortunately, that's not how it works.

A high-deductible plan takes people from bad shape to worse shape. Lastly, the researchers studied what deductibles did to people's health. Remember that families were randomized to the two health-insurance plans, so any differences in health outcomes during the RAND experiment years were probably a result of the impact of those plans.

After several years on the insurance plan that they were assigned to, everyone's health was evaluated. Overall, there was no difference: people who spent 3-5 years on the high-deductible plan finished the experiment in roughly the same health as people who spent that time on the free-care plan, according to a paper by Newhouse and a team of researchers.

Things were different, however, for one group of participants. The researchers focused on what they called "elevated-risk participants." That group consisted of people who were in poor health at the start of the experiment. Maybe they already had a chronic condition or maybe they were obese. For that group, the researchers found that the deductible plan harmed their health. A few years on a high-deductible plan took people from bad shape to worse shape.

That finding is, perhaps, intuitive. If you're in good health, a deductible will induce you to consume less healthcare, and that's going to have a very small impact on your health. After all, you're in good health, so a bit more or a bit less healthcare is not going to have a big effect, at least on average. But if you're already at elevated risk, a deductible leads you to consume less healthcare, and for you, that really matters.

The three conclusions of the RAND experiment paint a confusing picture of what deductibles do to people. On the one hand, deductibles lead people to cut back on healthcare a lot but they do not hurt people, on average. On the other hand, the participants who were assigned a high-deductible plan cut back on *all* healthcare, not just ineffective care. And the most vulnerable among them ended up worse off.—CBR

Tal Gross is a professor in the Department of Markets, Public Policy & Law at Boston University. Matthew J. Notowidigdo is the David McDaniel Keller Professor of Economics and Business and Public Policy Fellow at Chicago Booth. This is an edited excerpt from their book, Better Health Economics. *Reprinted with permission from the University of Chicago Press.*

Go to **chicagobooth.edu/review** to see citations for research mentioned in this article.

A new way to discover
research-driven insights
on business, policy, and markets.

Find us **every Wednesday** and subscribe
wherever you get your podcasts.

WOULD A NEW POLICY ON POT BE GOOD FOR THE US?

Marijuana may soon be getting a dramatic regulatory makeover in the United States. In May, the US Department of Justice submitted a proposal to reclassify marijuana from a Schedule I drug—a status it shares with heroin and LSD—to a Schedule III drug. The move wouldn't legalize marijuana for recreational use but would put it in the company of substances such as ketamine and anabolic steroids, which are considered to have moderate to low potential for dependency. It would also recognize marijuana as a drug with accepted medical applications and could affect the tax status of marijuana companies in states where pot is legal.

Would the effects of such a change be generally beneficial for society? To consider this question, Chicago Booth's Kent A. Clark Center for Global Markets polled its US panel of economic experts.

See more online
All responses to this poll can be seen at **kentclarkcenter.org**.

About the Clark Center Economic Experts Panels
To assess the extent to which economists agree or disagree on major public-policy issues, Booth's Kent A. Clark Center for Global Markets has assembled and regularly polls three diverse panels of expert economists, all senior faculty at the most elite research universities in the United States and Europe. The panels include Nobel laureates and John Bates Clark medalists, among others. Polls are emailed individually to the panel members, and panelists may consult whatever resources they like before answering. Members of the public are free to suggest questions.

PHOTO ILLUSTRATION BY EDMON DE HARO

Statement: Reclassifying marijuana as a Schedule III drug would lead to measurably higher social welfare.

Daron Acemoglu, MIT

"Legalizing marijuana in places where it's illegal, without interfering with state policies that have already legalized it, would be an improvement. There is no evidence that legalization has led to big negative effects, and further legalization would reduce criminal involvement."

Response: Agree

Judith Chevalier, Yale

"My understanding is that the compliance costs required for research labs to study the efficacy and harms of Schedule I drugs are much higher than for Schedule III. Lowering the barriers to such studies, given the widespread use of cannabis under state laws, seems a worthy policy goal."

Response: Agree

Erik Hurst, Chicago Booth

"Most states already allow medical marijuana use. So, in most states, it will have no effect. It may be best to decriminalize completely."

Response: Disagree

Fiona Scott Morton, Yale

"The relative dangers of these drugs are now well known. Marijuana is legal in many states."

Response: Agree

Nancy Stokey, University of Chicago

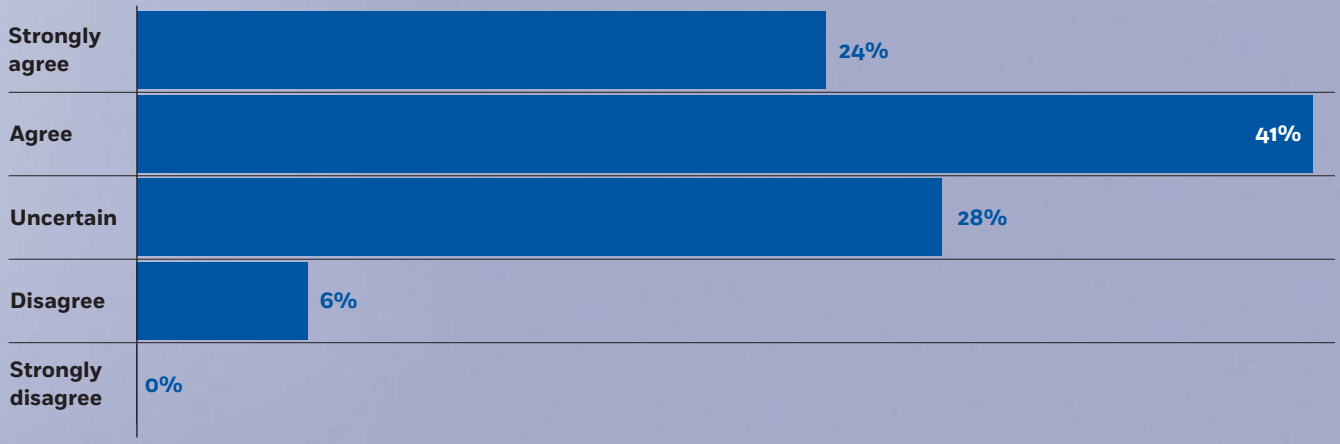
"Direct effects: lower costs for sellers through a tax change, and an upward shift in demand. Hence, sales will increase, leading to more and better targets for armed robberies, since it will still be a cash-only business. There will be winners and losers. On net, ambiguous."

Response: Uncertain

Chad Syverson, Chicago Booth

"Schedule I for marijuana is a decades-old mistake that created massive enforcement costs (direct and spillover). Moving away from it helps. Further moves in that direction may be optimal."

Response: Strongly agree



Note: Percentages are weighted by confidence ratings that panelists assigned to their own responses. Charts do not include panelists who reported "no opinion" or did not respond to the poll.

THE GRUMPY ECONOMIST
JOHN H. COCHRANE

I'm ready to be automated

How AI can shift supply and demand–perhaps with benefits for everyone

As I have been reading about and discussing large language models, I find I've learned as much about us humans as I have about the artificial intelligence that replicates some of what we do. Introspecting, am I really that much more than an LLM?

I recognize that I have about a thousand stories. Most of my conversations and writing, especially for my blog posts, op-eds, interviews, and discussions, are built on prompts that lead to those prepackaged stories. A given prompt could easily lead to a dozen different stories, so for a while I give the illusion of freshness to someone (not my wife and kids!) who hasn't been around me that long. *House prices are high in Palo Alto, should the government subsidize people to live here? Let me tell you about the vertical supply curve.*

Almost all of my stories are not original. I do a lot of reading and talking about public policy and economics, so I pick up more stories about those things than most people who have real jobs and pick up stories about something else. Learning and education are largely formal training for the acquisition of more stories to produce in response to prompts. That process is a lot like training a large language model.

This has got me thinking about programming a Grumpy Economist bot. Training an AI on the corpus of my blog, op-eds, teaching, and academic writing would probably give a darn good approximation to how I answer questions, because it's a darn good approximation to how I work.

I wouldn't be the first economist to be automated. George Mason University's David Beckworth, who hosts the *Macro Musings* podcast, has trained a Macro Musebot on more than 400 episodes of his show. Even Milton Friedman has been conjured algorithmically, courtesy of the Friedman chatbot at the University of Texas's Salem Center for Policy.

Now, not everything I do is complete recycling, predictable from my large body of ramblings or from what I've been "trained on." Every now and then, someone asks me a question I don't have a canned answer to. I have to think. I create a new story.

A great economist asked me for my intuition about how interest rates could raise inflation. It took a week to mull it over. I now have a good story, which helped me in writing a recent paper. Walking back with me to my office at the Hoover Institution after a seminar, Stanford's Robert Hall asked me how government bonds could have such low returns if they are a claim to surpluses, since surpluses, like dividends, are procyclical. The notion of an "s-shaped surplus process" and a whole chapter of my recent book, *The Fiscal Theory of the Price Level*, emerged after a few weeks of rumination. It's now a new story that I tell often. Perhaps too often for some of my colleagues.

This creativity seems like the human ability that AI will have a hard time replicating, though perhaps I'm deluding myself on just how original my new stories are. When I get that AI programmed up, I'll ask it the next puzzle that comes along.

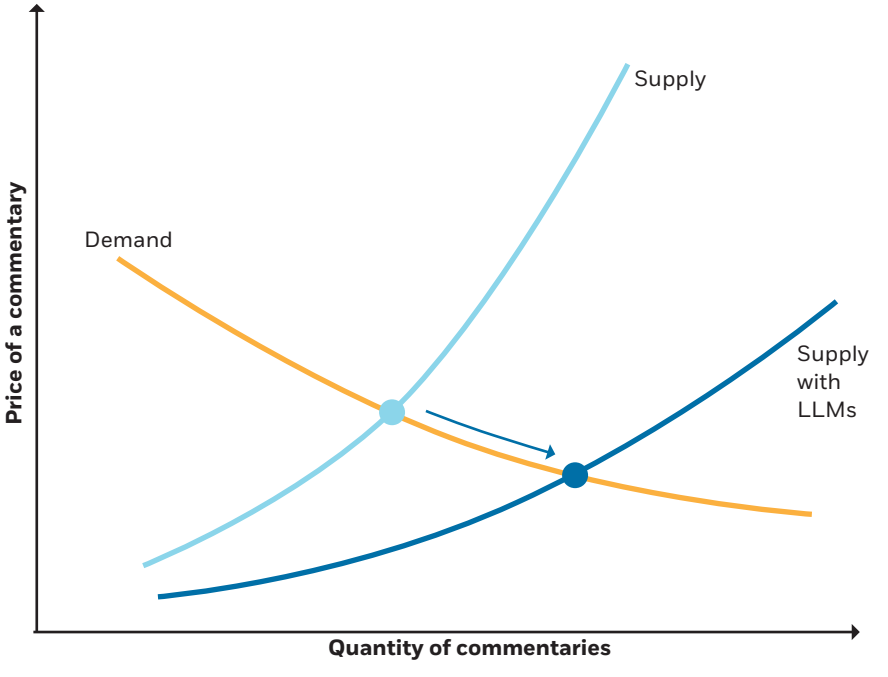
AI and the commentariat

This line of thinking leads me to recognize a part of my work that will certainly be greatly influenced by LLMs: the writing of blog posts and op-eds, the giving of interviews, and so forth. If 90 percent of what I do in that respect can be replicated, what does that mean for people in the commentary business?

Your natural instinct might be, "That business is toast and will be totally displaced by automation." Not so fast. Here is an old story, applied to this case. Look at supply and demand in the chart below:

Is a golden age of commentary approaching?

By lowering the cost of writing a blog post or op-ed, large language models could move the supply curve to the right and expand the demand for commentaries.



In the upper supply curve (rising to the right in light blue), I have the supply of commentary, along with where it intersects today's demand for commentary. LLMs push the supply curve down and to the right, as shown by the dark blue arrow and the new supply curve. I could certainly write more blog posts faster if I at least started with the bot and then edited. A colleague who is further ahead in this process reports that he routinely asks Claude.ai to summarize each academic paper in a 600-word op-ed, and he has found lately that he doesn't need to do any editing at all.

The curve shifts both down and to the right, however. We can produce more for the same total cost in time, or we can write the same amount faster.

Does that mean that the commentary business will end because the price will crash? Just asking the question in the context of supply and demand curves already tells you the answer is no. At a lower price, there is more demand, so the quantity expands. This could be the golden age of commentary. Indeed, quantity could expand so much that total revenue (price times quantity, or, in the chart, the size of a box with the origin in one corner and the supply-demand intersection in the other) could actually increase!

This has happened many times before. Movable type lowered the price of books. Did bookselling crash, and the monks starve? No. Demand at a lower price was so strong that bookselling took off, and more people made more money doing it. Though, as always, it was different people. The monks went on to other pursuits. Radio, TV, movies, and the internet each had the same effect on the communication industry. Technology that apparently substitutes for humans lowers costs, supply expands, and the market expands.

Automation and demand

It's not so obvious, though, that the demand for commentary is that flat. My inbox is already overwhelmed with papers colleagues have sent me to read and interesting-looking blog posts, and there are about 50 tabs open on my browser with more fascinating articles that I have not read. Related, the "price" in my graph, at least for this column, is the price of my time to produce it and the price of your time to read it. AI lowers the price for me, but not for you.

Now, what you need is an assistant who knows you and can read through all the mass of stuff that comes in and select and summarize the good stuff. That, too, is a task AI seems like it might be pretty good at. There's a joke (here comes another story I picked up somewhere) in which Joe says, "Look how great the AI is. I can input four bullet points and a whole PowerPoint presentation comes out!" Jane, getting the PowerPoint presentation, says, "Look how great the AI is. It boiled down this whole long PowerPoint into four bullet points!"

Of course, it has to somehow know which stories are going to resonate with you. Current algorithms are said to be pretty good, often too good, at feeding you what you like, but I want new things that expand my set of stories, and best of all, the rare things that successfully challenge and change my beliefs.

Indeed, perhaps AI will be more useful as digestion for information overflow than for producing even more to consume. I long wondered, what's the point of a lecture when you can just read the book? What's the point of a seminar when you can just read the paper? I think the answer is digestion. An hour-long lecture forces the professor to say what she can in that allotment. That's a short time, at best amounting to 10,000 words. Professors, at least in economics, notoriously assign endless reading lists that nobody could get through in a decade. In a lecture, they can't break the short time limit. They can lose everyone, or they can keep it digestible. Similarly, a good seminar with an engaged audience forces digestion.

In sum, perhaps AI will also help on the demand side, shifting demand to the right as well.

AI will need considerable oversight and hand-holding.

Implications for quality

Commentary is also a question of quality and not just quantity. Most commentary is pretty awful. Humans are not that good at reading critically, sticking to the point, maintaining logical continuity, avoiding pointless arguments, remembering basic facts, actually answering questions, and so on. At least the humans on my X stream aren't. AI editing might dramatically improve the quality of commentary. Just getting it from a C- to a B+ would be a great improvement.

As happens with all technology, AI will need considerable oversight and hand-holding. For the foreseeable future, there will be a need for humans to edit the output of the AI, figure out what prompts to give it to produce writing that will most interest readers, recommend and certify AI-produced material, and so on. The introduction of ATMs increased bank employment by making it easier to open bank branches and offer (overpriced) financial services. (You've probably heard that story. I've told it quite a few times.) Humans move to the high-value areas.

When I write a column like this one, I have to think things through, and often either the underlying story gets clearer or I realize it's wrong. If the AI writes all by itself, neither of us is going to get any better. But perhaps the editing part will be just as useful as my slow writing.

A good deal of what I learn from my work comes from conversations that my writing sparks with readers—online, by email, and in person—in which I often find my ideas were wrong or need revising. Once the comments are taken over by bots, I'm not sure that will continue to work. At least until I get a comment-reading bot going.–CBR

John H. Cochrane is a senior fellow of the Hoover Institution at Stanford University and was previously a professor of finance at Chicago Booth. This essay is adapted from a post on his blog, The Grumpy Economist.

IN-HOUSE ETHICIST
JOHN PAUL ROLLERT

The winner's dilemma in Liar's Poker

What a seminal book about Wall Street says about morals and moneymaking

This year marks the 35th anniversary of the publication of *Liar's Poker*, journalist Michael Lewis's first book and, by my lights, his most abiding.

Lewis was just 27 when he left his job at the investment bank Salomon Brothers at the beginning of 1988, the same age I was 17 years later when I first included *Liar's Poker* as the last entry on the syllabus for my business ethics class. A work that alternates between bildungsroman, brief history of the bond market, and a very '80s version of *Good-Bye to All That*, Lewis's book retains its power to prick the conscience, mostly by provoking uncomfortable laughter. And after the likes of Adam Smith, George Orwell, and Ayn Rand, the recovering bond salesman still gets the final say in my classes.

ILLUSTRATION BY MICHAEL BYERS

Today, the events of *Liar's Poker* are now closer in time to the Eisenhower administration than OpenAI, but the book is dated in ways that mostly don't matter. Yes, there is a striking affinity for pinstripes and flashy suspenders in its pages, but the promising young man we meet there still resembles most of my students. He's affable, enterprising, highly intelligent, and exceedingly well credentialed.

In other words, he's a winner.

Now, some of us may cringe at that distinction, Americans especially. The impulse toward social equality that Alexis de Tocqueville thought a hallmark of the national character—*We may have paupers aplenty, but no man's a king!*—keeps us from embracing such a label, at least unreservedly. The reticence makes for one of the more confounding tics of contemporary life. We are far more comfortable litigating the privilege that others enjoy than counting our own blessings accurately, a tendency that sees us routinely affirm an egalitarian commitment without ever pausing to consider the implications for our conduct and the sacrifices it might entail.

Lewis isn't nearly so inhibited. He knows that he's a winner.

But Lewis isn't nearly so inhibited. He knows that he's a winner. In fact, such a self-conception is central to *Liar's Poker*, for a growing ambivalence about that status is at the heart

of his story, nestled as it is in a world of high-stakes financial transactions where expediency and ethics seem radically divorced.

Getting one's just deserts

Lewis certainly feels like a winner when he arrives at Salomon Brothers in the fall of 1987. The moment is one of sweet vindication, for as a senior at Princeton, the art history major had failed to persuade any of the investment banks that visited campus to award him a callback. "I have never seen men on Wall Street in such complete agreement on any issue as they were on my application," he admits. "A few actually laughed at my résumé."

While making postgraduate amends at the London School of Economics, Lewis happens to be seated at a charity event next to the wife of a Salomon Brothers managing director. She takes a liking to him, and after a sustained grilling over dinner, she assures Lewis that "she would have her husband take care of it."

He does, and Lewis is welcomed to the Salomon Brothers training class of 1985. Its 127 members—chosen, Lewis learns, from more than 6,000 applicants—mostly hail from elite colleges and graduate degree programs. In this respect, Lewis is twice blessed, but he is sensitive to the fact that his immediate success is hardly a matter of strict meritocratic selection. "I decided to live with the stigma of having gotten my first real job through connections," he writes. "It was better than the stigma of unemployment."

Also better than unemployment is his first-year salary, which at $42,000 is nearly twice the US median household income. It's quite a sum for a 24-year-old, and on his way to his first day in the training program, Lewis is downright giddy. "I didn't really imagine I was going to work," he says of the occasion. It was "more as if I were going to collect lottery winnings."

"Lottery winnings" is an intriguing choice here. In my business ethics classes, I often use the example of a lottery to illustrate the idea of just deserts. There are many ways we can get the good things in life, and not infrequently, students who believe themselves to be defending free markets will instinctively resist any attempt to discriminate among them. If one is successful in getting the things one wants (so their reasoning goes), that's all that matters, and any attempt to make distinctions is a threat to a system that depends on the pursuit of self-interest.

The moral conundrum of *Liar's Poker* is what to do about unjust deserts when they're served up to you on a platter.

They may be right about this, but it doesn't mean that any such distinctions aren't keenly felt—or even that the students themselves don't already make them.

Take two hypothetical tales of rags to riches.

The first involves an engineering savant. We'll call him Boyle. Boyle comes from a working-class background—his father's a janitor at the local elementary school; his mother takes in laundry. Boyle likes school well enough, but going to college never crosses his mind. When his mother gets sick, he drops out of 10th grade and goes to work at a local garage to help the family.

He enjoys tinkering, and he's good with his hands. After a few years as a mechanic, he notices that the steel fasteners on the panels of the pick-up trucks that come into the shop keep rotting out. They're hard to remove and slow down the work, which always leads to a lot of complaints. One day, Boyle gets an idea for replacing them. He drains his savings, cuts back to one beer on bowling night, and sells the antique pocket watch his grandfather left him. It's just enough for him to buy a small injection-molding machine, which he sets up in his basement.

He starts spending nights and weekends creating a plastic clip to replace the steel fasteners, one that can be easily removed and will better weather the elements. After two years of testing, he shares his prototype with a parts supplier for Ford and persuades it to give him a shot. It's a success. Boyle buys back his grandfather's watch, and within a decade, he's the owner of a small plastics company that employs almost 20 people and earns nearly $2 million a year.

The second story involves another individual. We'll call him Barneby. On the way home from softball practice, Barneby stops at a gas station to refill his Audi. On a lark, he buys a Powerball ticket and wins $250 million.

Now, consider both men—is there any way they're equally deserving of their riches? Of course not. In fact, to say that Barneby *deserves* his lottery winnings seems like a confusion of terms, one that conflates sheer luck with simple justice. In contrast to Boyle's entrepreneurial odyssey, Barneby didn't *do* anything to become a centimillionaire—nothing, at least, beyond buying a lottery ticket.

When the behavior required is deplorable, it's helpful to nourish an effortless scorn for those disadvantaged by your endeavors.

Customers' short memories

Michael Lewis certainly sees dumb luck at work in his excessive wages, but he isn't terribly bothered by the inequity. Life is hardly fair, and between his good humor and congenital insouciance, he's not the type to be afflicted by any great sense of guilt for a little good fortune. "There was no justice in the world," he cheekily notes, "and thank goodness for that."

Besides, Lewis is among the ranks of the best and brightest. Surely whatever they do is worth a prince's ransom.

The sorry truth makes for some of the most memorable moments in *Liar's Poker*. Readers are soon introduced to the sordid business of being a bond salesman, an aggressively vulgar world of "jamming bonds" (trading floor parlance for persuading clients to buy dodgy securities) and "blowing up customers" (what happened when they did so). Lewis does a bit of both during his two years in investment banking, an experience that doesn't sit easily with him. "I'm now convinced," he writes, "that the worst thing a man can do with a telephone without breaking the law is to call someone he doesn't know and try to sell that person something he doesn't want."

Like the scenes of a slasher film, the stories Lewis shares are lurid and gripping, and one might be forgiven for wondering how exactly Salomon Brothers ever hoped to operate for very long treating its customers like garbage. Lewis wonders too. "A policy of screwing investors could lead to ruin," he writes. "If they ever caught on, we'd have no investors. Without investors, we'd have no business raising money."

The book offers two explanations. One involves the type of customers a junior salesman tended to draw, "small clients"—those who "if disaster struck" and a trade blew them up, "the effect on [the] overall business of Salomon Brothers would be negligible." The other is courtesy of the firm's president, Tom Strauss, himself a former bond salesman: "Customers have very short memories."

Strauss's remark is foul, but in Lewis's telling, the sentiment supporting it is hardly exceptional. A vicious contempt for customers seasons the boorish banter of the trading floor and is central to the survival-of-the-fittest ethic that seemingly guides everyone's behavior. For instance, early on at Salomon, Lewis is duped by a trader into helping unwind a bad position by selling $3 million of bonds to one of his clients, a trade that results in an immediate loss for the investor. Guilt-ridden and furious, Lewis knows the "best thing" he can do is to "pretend" that he had "meant to screw the customer." Why so? "People would respect that."

But they wouldn't just respect him for his decision; they would pay him handsomely for the deceit. At the end of his first year, Lewis earns a $45,000 bonus, making him among the four highest-paid members of his training class and more than doubling his salary.

Escaping the lion's den

If the difference between the tales of Boyle's plastic fastener and Barneby's lottery ticket is one of simple and just deserts, the moral conundrum of *Liar's Poker* is what to do about unjust deserts when they're served up to you on a platter. "It was more than I had contributed to society," Lewis says of his full first-year compensation. "Christ, if social contribution had been the measure, I should have been billed rather than paid at the end of the year."

The irony only grows with Lewis's continued success. "My father's generation grew up with certain beliefs," Lewis writes after his second-year compensation swells to $225,000. "One of those beliefs is that the amount of money one earns is a rough guide to one's contributions to the welfare and prosperity of our society."

Lewis clearly shares this belief, and the question for him is what to do when you're presented with clear evidence to the contrary, when you find yourself playing some "absurd money game" and benefiting "out of all proportion to your value to society."

This is the winner's dilemma. It's a conundrum that many of my students face, a monkey's paw for clambering to the very top of the meritocratic heap. By the remorseless metrics of the modern educational system, they have proven themselves elite athletes of the mind, individuals capable of organizing, analyzing, and synthesizing information, of conceiving projects and leading teams. And for all of their troubles, if they succeed, they're offered a chance to make lavish sums of money for doing things whose value to society is highly questionable if not outright injurious.

Set aside the broader implications for any community that channels its best and brightest into such infelicitous busywork. The very personal danger, Lewis discovers, is that to justify such choices, especially when the behavior required is deplorable, it's helpful to nourish an effortless scorn for those disadvantaged by your endeavors. The reason is simple: if winning is hateful, it's easier if the losers deserve it. Lewis recounts a moment early on at Salomon Brothers when a hapless associate gets swatted aside by a superior on the trading floor and visibly begins to panic. "What a wimp," Lewis says he instantly thought, then he realized of his reaction, "it showed I was coming along."

The moment captures the twisted meritocracy of Salomon's trading floor, whose moral logic Lewis describes as "capitalism at its most raw" and "self-destructive." Whenever I return to it, the passage puts me in mind of something I always tell my students, an existential lesson of sorts: however sturdy and self-possessed you think you are, there is not a hard kernel of you-ness that persists over time like some indestructible diamond. We are not the sum total of our experiences, but environments do tend to change us, and we are all subject to the logic of the lion's den. Enter one, and you will quickly learn to bite or you will get bitten.

It took guts, real guts, for a 27-year-old to look at his swollen bank statement and say: *To hell with this.*

Either way, you'll eventually be a beast.

That transformation is not immediate, however. It takes time. After two years at Salomon Brothers, Lewis admits that, when he looked in the mirror, he didn't recognize the person he saw staring back. Some day he would, though, and then the change would be complete.

And so, before there was no turning back, Lewis describes making "as stupid a financial decision as I hope I'll ever make." He collected his second-year bonus and quit.

Lewis was convinced he was walking away from the "clearest shot" he would ever have at becoming a millionaire. That may have been true at the time, but his readers know he turned out all right in the end. Today Lewis is not only one of the most successful authors of his generation, but he's also a millionaire—many times over.

Still, that doesn't take away from the audacity of his decision to leave Salomon. A million dollars is an enchanting number at any age, and Lewis occupied a world where numbers like this mattered. Dollars were a way of keeping score—the *only* way of keeping score—so it took guts, real guts, for a 27-year-old to look at his swollen bank statement and say: *To hell with this. I've got a fantastic education, opportunities galore, and a lifetime ahead of me. This is a zero sum game—I'm better than this.*

Such courage is ultimately the reason I've assigned *Liar's Poker* to so many students over so many years. There's being a winner and then there's acting like one. The lesson of Lewis's book is the price of knowing the difference.

"You might say that I left the trading floor of Salomon Brothers in search of risk," Lewis says in the closing pages of the book. He decided to bet on himself, on his highest sense of self.

Thirty-five years later, it's still a thrilling gamble—and one too few winners make.—CBR

John Paul Rollert is adjunct associate professor of behavioral science at Chicago Booth.

How to improve a city commute

Many cities would like to ease road congestion, but their leaders can't agree how best to do so. Making public transit rides cheaper and more frequent would require funding, while instituting road taxes could be unpopular. But combining these policies—using revenues raised from road taxes to subsidize public transportation—could ease the burden on middle-income commuters, suggests a study by a group of researchers that includes Chicago Booth's Milena Almagro. The researchers analyzed the impact of various transport policies on Chicago commuters, focusing on consumer surplus—the benefit derived from a transportation mode under a given policy minus the cost to commuters of using that mode. By aggregating across different modes (such as public transit, private car, or ride-hailing service) and consumer types (grouped according to income, transport preference, and car ownership), the researchers were able to measure the surplus generated by different policies and combinations of them. To learn more, turn to page 22.

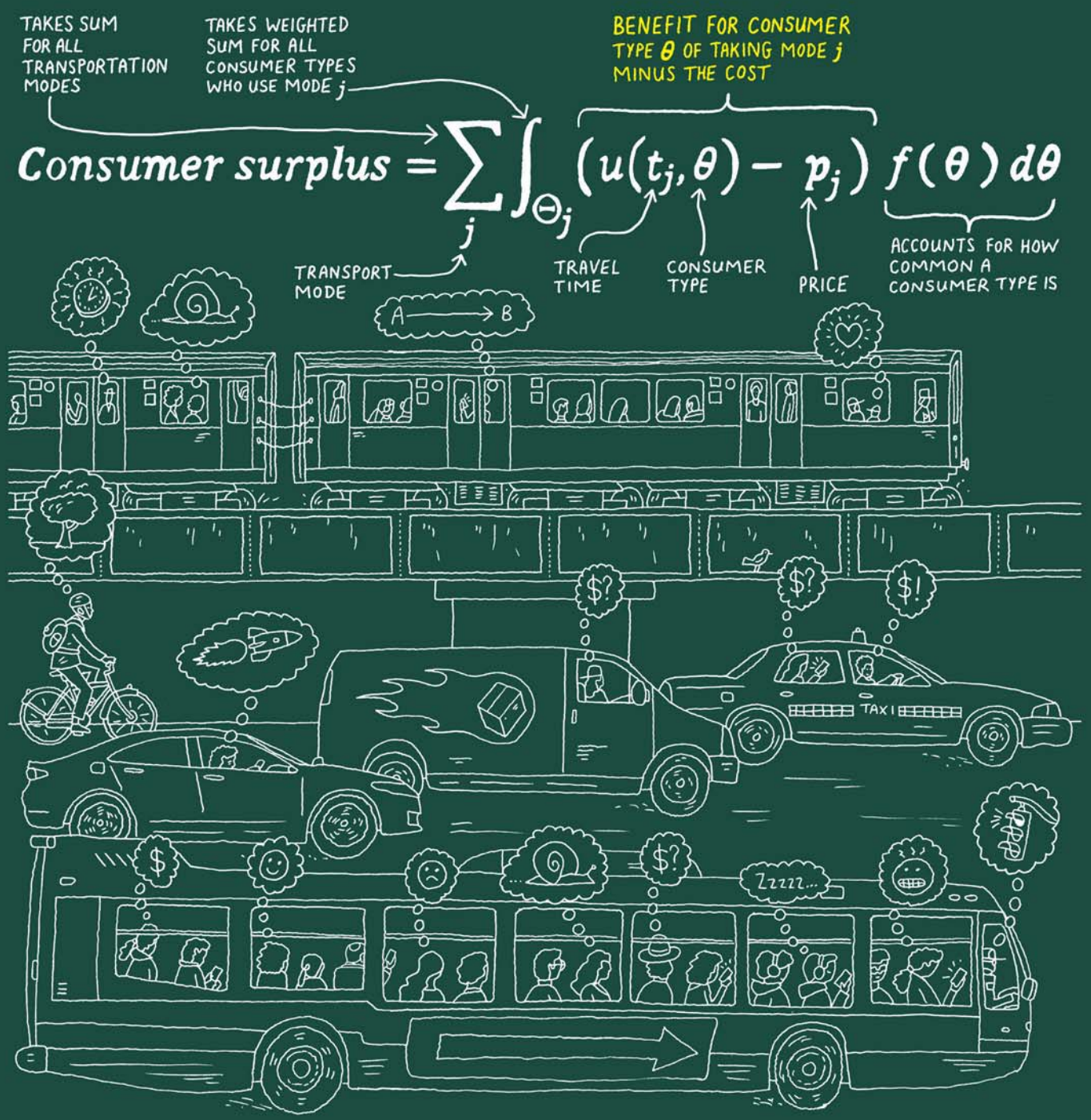


ILLUSTRATION BY PETER ARKLE

CHICAGO BOOTH
The University of Chicago Booth School of Business
Chicago Booth Review

The University of Chicago
Booth School of Business
5807 S. Woodlawn Ave.
Chicago, IL 60637

HOW TO GET PEOPLE TO ACCURATELY PREDICT PRICE CHANGES
Page 18