Saurabh Bhargava

Summary of Research and Scholarly Contribution

Department of Social and Decision Sciences
Carnegie Mellon University

PLEASE CLICK HERE FOR UPDATED STATEMENT & MATERIALS

1. RESEARCH OVERVIEW

As an applied microeconomist, specializing in Behavioral Economics and Public Policy, my research broadly explores how insights into the cognitive, social-psychological, and emotional foundations of decisions can enrich understanding of real-life behavior—particularly among at-risk populations—in settings of economic importance. Part of this agenda is to develop the implications of this psychologically-informed understanding for improving the design, implementation, and marketing of policies and programs, as well as the descriptive accuracy of models used by economists for policy and welfare analysis. Methodologically, my research relies on an array of identification strategies—such as large-scale field experiments, natural experiments, and analysis of high-frequency data—often leveraging access to exceptionally unique data or partnerships with government agencies or large firms. I almost always supplement statistical insights from these analyses with data from hypothetical choice experiments, baseline surveys of beliefs and decision processes, administrative data, and expert interviews to better understand baseline incidence, disentangle psychological mechanisms, explore heterogeneity, and address possible confounds.

My research is distinguished by its focus on empirical puzzles of central importance to theory or welfare, sometimes explored through ambitious, multi-year, investigations. I've organized the discussion of my work within three broad topic areas, each motivated by a series of focal puzzles. In the first section, Take-Up of Benefit Programs, I describe research attempting to address a long-standing puzzle as to why individuals do not claim valuable government benefits for which they are eligible and whether one can interpret incomplete take-up as the use of ordeals to achieve optimal targeting. In the second section, Risk and Insurance, I discuss research that leverages a natural experiment to examine whether insurance choice can be interpreted from the lens of standard economic models; analyzes a unique employee-rewards program to generate new insight into the motives of financial risk-aversion (including a novel heuristic explanation); and links driver telematics and claims data from a large insurer to surveys of the same drivers' risk forecasts and behavioral biases to propose a new explanation for the puzzling lack of adverse selection observed in the insurance literature. In the third section, Retirement Saving in the U.S., I describe research examining candidate explanations for several demand-side saving puzzles and proposing a potentially unifying explanation. I also describe research identifying a supply-side friction in the design of 401(k) automatic enrollment that, if addressed, could potentially lead to a greater increase in retirement preparedness than the introduction of automatic enrollment (this friction also has significant implications for consumer protection and fiduciary duty). Papers in this section collectively describe findings from field experiments across 899 firms, hypothetical enrollment experiments involving an additional several thousand employees, and surveys of dozens of high-ranking policymakers and several hundred plan executives.

In the final section, I discuss several extensions, including a series of papers on Emotion and Well-being and papers Translating Effects from Lab to Field. The first category includes papers that analyze an experience sample with unprecedented detail on emotion, mood, and time-use to resolve a long-standing puzzle of kids and parental happiness and to provide an empirical account of the experience of love. A paper in the second category exploits a natural experiment—the discontinuous rise in cell phone use at 9pm caused by the switch from peak to off-peak pricing— and nearly two dozen datasets to provide new evidence on the (pre-smartphone era) link between driver cell phone use and crash risk. While the paper relied on several proprietary datasets on crashes, pricing plans, and call data, it likely featured the first academic use, in the social sciences, of large-scale data on cell phone signals and their routing patterns across towers (we used these data to infer cellular use among drivers). Our findings, rejecting a consensus of 120+ studies, were corroborated by RCTs subsequently administered by the National Highway Traffic Safety Administration.

1.1 Scholarly Impact and Contributions

Overall, I see my work as having had a significant influence on the field, policy formation, and policy outcomes. In addition to a more specific discussion below, I briefly summarize this influence here. While two of my

most important papers (BC 2022; BCMB 2021), are currently under review, as evidenced by various markers, I believe that the field has seen my earlier work as seminal. For instance, Behavioral Economics has an annual meeting (BEAM) in which roughly 10 to 12 papers are selected for presentation to a plenary session of leading researchers worldwide. Three of my papers have been presented in these meetings (BM 2015; BLS 2017; BC 2022). My work has also been generously represented in prominent summaries of the field such as, Raj Chetty's 2015 *Ely Lecture on Behavioral Economics and Policy* (BM 2015; BLS 2017), the Handbook of Behavioral Economics (BF 2014; BM 2015; BLS 2017), and a report for the National Academy of Sciences' Committee on Behavioral Economics and Policy (BM 2015; BLS 2017; BC 2022). One paper won an award from the BSPA for the Best Publication of the Year and was a finalist for the prestigious NIHCM's Annual Health Care Research Award (BLS 2017). Finally, three papers (BM 2015; BL 2015; BLS 2017) are often included in syllabi for courses in Behavioral Economics or Public Finance at leading institutions; less frequently two other papers are included (BF 2015; BL 2015).

Methodology. I'd suggest my distinctive strengths as a researcher are the ability to access exceptional data or forge unique institutional partnerships as well as the imaginative application of methods to empirical problems. A primary example is the use of a field experiment, conducted in 2011, to study incomplete take-up of social benefit programs (BM 2015). To my knowledge, this was the first study to investigate the causes of incomplete social benefit take-up with a field experiment, the first field experiment emerging from an academic collaboration with the IRS, the first field experimental test of stigma on program take-up, and presumably one of the earliest behavioral science field experiments with the US government, predating the SBST by 4 years. Subsequent methods I've adopted were motivated by the concern as to how to credibly draw population-level inferences from low-compliance settings (e.g., experiments using emails, notices, texts where response can be very low and unrepresentative). In the 2011 EITC experiment, I addressed the compliance problem by supplementing field experimental evidence on the importance of "psychological frictions" at the margin with psychometric analyses of the experimental treatments and surveys of thousands of low-income tax filers to document the baseline incidence of such frictions (an approach I had not observed in the prior literature). The same concern over low-compliance and ecological validity led my coauthor and I, in our study of retirement savings (BC 2022) to embed a detailed survey capturing the baseline incidence of various psychological friction within a field experiment experimentally manipulated those frictions for the same employee sample. This strategy permitted us to estimate the baseline prevalence of each friction (and its naïve correlation with saving), the average effect of reducing a specific friction on saving, and the differential effect of reducing a specific friction on saving across baseline incidence. In newer research, considerations of compliance shape the choice of data setting and the implementation of field experiments (BC 2022; BH 2022; BCMB 2021). For example, in a field examination of the motives for financial risk-taking, we analyze choices (and beliefs) from more than 20,000 employees participating in an employee-rewards program with 98 percent participation (BH 2022).

As another example, I have been increasingly drawn towards the using "virtual" decision-environments as an inexpensive strategy for pre-testing large-scale field interventions or even policy proposals. The strategy isn't simply an endorsement of experimental pre-tests. Instead, it's about identifying critical decision-relevant features of a complicated real-world setting (such as insurance, or 401(k) enrollment), incorporating such features into stylized lab environments, and then validating those environments with field data through a lab-field correspondence in baseline choice, price/cost sensitivity, and even response to simple interventions. I have applied this approach to better understand decisions in insurance, saving, patient decision-making, and even response to newly proposed food labels. Other examples of novel application include borrowing techniques from labor economics (Oaxaca-decomposition techniques, propensity scores) to estimate gender differences in love (BK 2022) and the causal effect of children on well-being (Bhargava 2022), the use of simulations and annuity-pricing to characterize retirement risk by the probabilistic risk of retirement poverty, an arguably more welfare-relevant outcome than the average ratios of income replacement typically used in the literature.

Theory. Finally, while most of my work seeks to understand whether empirical phenomena (such as take-up, insurance choice, financial risk-taking, retail investing) can be understood through existing (behavioral) economic frameworks, some of my work has advanced new theoretical explanations for such phenomena. A first example is a hedonic, belief-based, model of present focus presented in BC 2022. The framework offers an alternative to more popular preference-based explanations for delayed action and a lens from which to understand response to small incentives (or rewards) and anxiety, informed by research on the neural mechanisms underlying impulsivity. While I hope to explicitly test and develop the model in future research, the paper describes how the model could plausibly

account for a collection of savings puzzles (BC 2022). As a second example, in a new paper clarifying the motives of financial risk-taking, we propose and test a novel decision heuristic that appears to explain risk-averse financial choice from certain menus more accurately than standard economic and behavioral benchmarks (BH 2022). From a decision theory perspective, I highlight two novel cognitive (or motivational) phenomena from my research that I plan to explore further and summarize for a psychology or decision research outlet. The first is a systematic empirical pattern of delayed optimism about future hedonic relief. This insight underpins the aforementioned model of belief-based present focus and has not, to my knowledge, been discussed in the literature on optimism bias or hedonic forecasting. The second is the appreciable underestimation of correlated pairwise probabilities that we repeatedly document in BH 2022.

Policy. I take considerable pride in my work's direct (and direct) influence on policy/program design and individual welfare. I think it would be reasonable to say that the direct impact of my research on at-risk individuals in the US is comparable to any Economist in my cohort. I describe several examples: (1) During my research on program take-up, I redesigned several EITC tax notices and claiming worksheets subsequently adopted and used for several years by the IRS—we estimate that these redesigned forms would lead to an additional 800k claimants of $504m in benefits each year (the research prompted benefit-program redesign efforts in other nations). An unpublished experiment I designed and administered resulted in redesigned tax notices the agency adopted for EITC Tax Preparers. (2) Research on insurance choice led to conversations with the CMS about the design and marketing of the ACA exchanges and prompted a 2017 workshop where industry representatives convened to discuss our research's implications for future insurance product design. More directly, one of our partner firms responded to our findings by reforming the health plans offered to their 51k benefit-eligible employees (thereby undoing the negative welfare transfer we had identified). (3) Research on the supply-side dynamics of retirement savings, led us to identify a potential friction in the auto-enrollment process that, if corrected, could have a larger projected impact on savings/preparedness than the celebrated introduction of automatic enrollment. This work also led to redesigned enrollment interfaces now engaged by several hundred thousand potential enrollees each year (a reform whose estimated effect on savings is equivalent to a 68 to 74 percent increase in the match). I expect an insight from current work on auto-escalation to, if adopted, lead to a significant increase in accumulation for a few million (or more) enrollees at our partner financial service provider. (4) Research on the demand-side barriers to retirement savings led us to identify a new potentially first-order source of under-saving (confusion about enrollment status) and a novel strategy for increasing engagement among at-risk employees. I have begun to engage industry and policy contacts on both issues. The same research led me to design a new model for a retirement savings plan (The Serenity Plan) that I have also began circulating.

I invest substantial time engaging audiences beyond academia to amplify my research's potential impact. For example, I have advised several government agencies, behavioral science ("nudge") units, and social policy organizations including the Internal Revenue Service, Department of Justice, Consumer Financial Protection Bureau, the US Securities and Exchange Commission, MDRC, Allegheny County Pretrial Services, Centers for Medicare & Medicaid Services, Ideas42, Elizabeth Warren's Senate Campaign, the United Way, and the Pennsylvania Insurance Department. I've also participated in expert roundtables and/or forums hosted by The Brookings Institution, CDC, GAO, NIH, Laura and John Arnold Foundation, the Russell Sage Foundation, and the Aspen Institute. As a part of my work with the Aspen Institute on issues pertaining to retirement security, I moderated one panel featuring senior industry executives and an expert from McKinsey & Co., participated on three additional panels as an expert, and am one of a handful of US academics invited to their annual leadership retreat on retirement reform. Recently I agreed to join an exploratory Aspen working-group tasked with developing the implications of Blockchain/Defi for households.

As another example, when, a few years ago, the Behavioral Science Policy Association assembled eight academic teams, each comprised of ten to fifteen leading scholars, to draft policy blueprints for the White House SBST (the US nudge unit), I was the only researcher in the country to serve on two policy teams. I contributed to a blueprint for health policy reform (Loewenstein et al. 2017) and a blueprint reform of policies related to household finance (Madrian et al. 2017). I was honored to have been one of eleven people to present policy recommendations, earlier this year, to a National Academy of Science panel on Behavioral Science (a group that included Richard Thaler). Finally, I have informally advised several municipal or state/federal nudge units, was once invited to join the White House's SBST, and in 2016, was asked to lead Qatar's new nudge unit (a flattering offer I politely declined). I was selected as an affiliate of the Jameel Poverty Action Lab at MIT (J-PAL) in 2018, currently the only CMU affiliate, and a Visiting Scholar at the Russell Sage Foundation in 2017.

1.2 Note on Health Context

I experienced a challenging health stretch in 2016 that negatively affected my professional and non-professional life. I have been advised to discuss this briefly. Around that time, I was diagnosed with testicular cancer, and surgically treated at Memorial Sloan-Kettering. Early-stage seminoma, of the type I contracted, typically has excellent prognoses (I recently completed my fifth year of surveillance). While patients often experience a post-surgical hormonal imbalance—sometimes severe— it often subsides within months. Unfortunately, in my case, I experienced a severe imbalance lasting over a year (at which point my Testosterone levels registered as average—for anyone 75+ years-old—a situation jovially described by one urologist as a "medical anomaly"). While I have since adjusted to a more moderate imbalance, the practical outcome is that I regularly experience long periods of fatigue, limited attention, diminished metabolism, and a loss of ability to sustain multi-day stretches of intense effort when needed. I have been on a reasonably stable mix of treatments recently and anticipate switching to a new, hopefully curative, regimen within two years (one I long avoided due to fertility risk). My medical situation interacted negatively with the pandemic, a circumstance I recognize was a hardship for many. With all this shared, I believe my productivity and the pacing of my research, while compromised for a stretch (and requiring medical leave), has returned to an encouraging place. On a merrier note, my medical misadventures led to productive introspection. For example, my experiences directly inspired the hedonic model of belief-based present focus described above. It also inspired experiments investigating the effects of integral label-induced anxiety on patient decisions (e.g., whether referring to a condition as low-grade cancer versus a neoplastic lesion, all else equal, affects patient risk-taking) and an empirical investigation as to whether observed variation in clinical treatments reflects clinician heterogeneity as opposed to shared patient-clinician decision-making.

2. RESEARCH BY TOPIC AREA

2.1 Take-Up of Benefit Programs

A well-documented, and perhaps surprising, feature of social benefit programs in the US is that many eligible individuals fail to claim their benefits. The consequences of this incomplete take-up can be significant. In the EITC, the nation's largest means-tested cash transfer program, an estimated 6.7 million, or 25 percent of those eligible, fail to claim an equivalent of 33 days of average income (Plueger 2009). While increasing social benefit take-up in programs such as the EITC is an unequivocal objective, the rationale is less obvious from the perspective of the standard economic model. If existing barriers to claiming a credit, such as the time/effort-costs of claiming, deter applications from those with little economic need, such ordeals may serve as an efficient sorting mechanism. If instead non-claiming is due to "psychological frictions"—such as confusion about program eligibility or benefits or psychological aversion to complexity—increasing take-up could be welfare improving, significantly so if such frictions disproportionately affect those with in greatest economic need. If such frictions did exist, they would signal a rejection of the standard economic model of take-up and of policy design/welfare analyses informed by the model.

To explore the causes of incomplete benefit take-up, and in response to a call by Curie (2006) noting the lack of causal evidence on social benefit program take-up, I conducted a large policy experiment in 2011 in collaboration with the EITC Office (Bhargava and Manoli, Psychological Frictions and Take-Up, *American Economic Review* 2015, BM 2015). The field experiment randomized the distribution of strategically modified claiming worksheets and notices to the universe of 35,050 tax filers from California who failed to claim their 2009 TY EITC despite presumed eligibility and the receipt of a first reminder notice. The EITC offers a favorable setting to investigate low take-up since the conventional costs of claiming are low (requiring the completion and return of a 1 to 2 page claiming worksheet) relative to the potentially substantial benefits of claiming. For all recipients, the mailings (a reminder notice, claiming worksheet, and a return envelope) conveyed program eligibility and provided recipients an opportunity to claim. The mailings varied experimentally in the content and complexity of the notice and worksheet, the presence of an additional informational flyer, and messaging on the envelope. The differential response to the mailings permitted inferences about the relative importance of three potential explanations for non-claiming: *program confusion* about eligibility or benefit size, the time-required to enroll or the likelihood of an audit, the *complexity* of the notice and/or claiming worksheet, and the perceived (social and personal) *stigma* associated with the program. Finally, for insight into mechanisms and to understand the baseline prevalence of psychological frictions across a broader sample of non-claimants, we supplemented the field experiment with in-person surveys of 1,139 low-income taxpayers at volunteer tax-clinics and a

psychometric assessment of the experimental interventions involving 2,800 online participants. Collectively, the intent of the investigation was to generate a causal understanding of the broad causes of incomplete take-up. All told, the experiment informed individuals of $26 million in unclaimed benefits, of which $4 million was claimed.

Findings. A first finding is to document the sensitivity of benefit take-up to non-economic variation in the frequency, salience, and simplicity with which benefit information was provided. Notably, merely receiving a second opportunity to take-up—that is, simply being included in our experiment—led 0.22 of the sample to claim. Across interventions, simplifying mailings (either through visual design or a shorter worksheet excluding non-diagnostic criteria) and displaying potential benefits in the headline collectively more than doubled take-up relative to baseline (0.14 to 0.31, $p < 0.01$). A second finding is to implicate select psychological frictions—lack of program awareness and confusion about eligibility and/or benefit size— as causes of incomplete take-up. We reach this conclusion from surveys indicating widespread lack of EITC awareness and self-knowledge of eligibility and among the eligible, widespread, and substantial underestimation of benefit size and psychometric assessments indicating that the interventions most successful in increasing take-up, increased program awareness, perceived eligibility, and perceived benefit size. A third finding is that confusion regarding the time-costs or the likelihood of an audit did not appear to contribute to low take-up—potential claimants had accurate perceptions of the time required to enroll and experimental messaging to reduce this perception did not increase claiming and while they overestimated audit likelihood by 15x, experimentally reducing such perceptions did not increase claiming. A fourth finding is to reject the role of stigma (either personal or social) in incomplete take-up— we did not find high baseline perceptions of stigma and interventions reducing stigma did not increase claiming. A final finding is that, among several patterns of heterogeneity reported in the paper, remedying program complexity disproportionately raised take-up among low earners and those with dependents.

Contributions and Impact. I see this paper as my most influential and innovative to date (while coauthored, I conceptualized the research, designed the experiment and surveys, and drafted the paper). Methodologically, to my knowledge the study, conducted in 2011, is the first attempt to investigate the psychological and economic causes of incomplete take-up in a government benefit program via a field experiment (the idea was inspired after I read Bertrand et al. 2010's innovative experimental examination of non-economic cues in the demand for loans in South Africa). It is the first experiment emerging from an academic collaboration with the IRS and presumably one of the earliest behavioral science field experiments within the US government (the study predated the SBST by 4 years). Due to concerns over how to draw population-based inferences as to causes of incomplete take-up, I combined evidence as to the existence of frictions on the margin from the field experiment with psychometric assessments as to the effect of treatments on beliefs and decision-making and with surveys documenting the baseline incidence of frictions across a broad population of eligible individuals. These surveys involved several months volunteering at VITA tax centers to oversee survey distribution and to speak individually to hundreds of low-income filers. The paper continues to be one of the few field tests of stigma in benefit take-up (Schanzenbach (2009) notably tested how destigmatizing SNAP affected interest in learning about the program).

Theoretically, the paper offered perhaps the strongest evidence that incomplete benefit take-up is a policy problem that reflect psychological frictions, disproportionately affecting those most in need, rather than reflecting the use of administrative ordeals to optimally target benefits. It also provided a direct challenge to standard economic models of take-up, contributed to then-modest literature documenting the role of information provision and simplification on benefit program engagement (Hastings and Weinstein 2008, Kling et al. 2012; Beshears et al. 2013), contributed to the literature documenting a lack of consumer literacy in benefit programs (e.g., Chetty et al. 2013), and provided the first evidence for systematic underestimation of eligibility/benefits of which I know. Finally, it was an early demonstration as to the regressive welfare effects of complexity (following Dynarski and Scott-Clayton 2006).

Practically, the paper illustrated a low-cost, scalable, strategy through which to improve social benefit take-up among a population not highly responsive to changes in benefit size. In the paper we estimated that expanding the population of recipients, redesigning documents, and instituting a second mailing to initial non-respondents, could improve EITC take-up from 0.75 to 0.83—of which 0.03 (approximately 800k claimants and $503m in benefits) was attributable to the redesigned mailings we tested. While the EITC as of 2021 no longer used the redesigned CP notices, they were used for several years after the initial study and prompted subsequent EITC efforts to increase program marketing. The research also prompted benefit program redesign efforts in countries outside the US. The paper was presented at BEAM in 2012.

2.2 Risk and Insurance

Economists—and social scientists more broadly—have long sought to understand the motives for financial risk taking. Clarifying such motives has profound implications for economic theory, the measurement of consumer welfare, and the optimal design of programs and policy. From the perspective of standard economic theory, risk-aversion among fully-informed, utility-maximizing individuals should reflect the diminishing marginal utility of wealth (von Neumann and Morgenstern 1947). Existing research, from the lab (e.g., studies of hypothetical gambles) and the field (primarily inferences of risk from insurance demand or investing behavior), however, has challenged the expected-utility (EU) framework. In particular, people exhibit a degree of risk aversion that is greater than that predicted by utility-based preferences for risk and empirical variation in perceived risk and risk attitudes explains a limited amount of observed risk-taking. Given the fundamental importance of the question, researchers have advanced a number of modifications to the utility-based framework to explain risk aversion through alternative channels such as biased beliefs, non-linear decision weights, or loss (or disappointment) aversion in the context of gain-loss utility (see Dhami 2019). Such aversion to risk could also emerge from psychological processes that have received less attention in economics such as those involving heuristics, salience, affect, cognitive processes, or hormones.

In this section, I discuss a series of projects that leverage multiple unique data environments, along with several online experimental paradigms, to: (1) answer whether insurance demand can be interpreted through utility-based models of risk, (2) investigate whether financial risk-taking in the context of a novel all-or-nothing employee-rewards program (characterized by simplicity, generalizability, and rich data on beliefs can be explained through existing economic models, (3) propose a new heuristic explanation for risk-averse behavior from menus, and (4) develop the implications of these findings for the design of insurance exchanges and as potential solutions for empirical puzzles in the insurance literature.

2.2.1 Demand for Dominated Insurance Plans

There is an active debate in the economics literature over what the demand for deductible-insurance implies about attitudes towards financial risk and how to use insurance-based estimates of risk to inform welfare analyses. A challenge for those evaluating insurance choice, however, is that, such choices could, in theory, reflect a range of financial (risk, risk preferences, wealth) and non-financial (e.g., network coverage, insurer reputation) considerations often unobserved. My colleagues, George Loewenstein, Justin Syndor, and I overcame these challenges by analyzing a setting that offered unprecedented clarity as to whether one could explain insurance choice through standard economic models (**Bhargava**, Loewenstein, Sydnor, 2017, *Quarterly Journal of Economics*, BLS 2017).

The exceptionalism of the setting owes to the decision of a large US firm to permit employees to fully customize their health insurance. This was accomplished by employees "building" a plan from a menu with effectively 48 options by sequentially selecting their preferred level of each of four cost-sharing plan components. For example, employees could select a preferred deductible ($350, $500, $750, $1,000) before proceeding to select a copayment, coinsurance, and out-of-pocket maximum. Three features of the menu inadvertently created an ideal litmus test for researchers to investigate choice efficiency. First, besides differences in cost-sharing and cost, the 48 plans were otherwise identical. Second, unlike major exchanges like Medicare Part D, the enrollment interface was simple and highly standardized. Finally, because of plan pricing, a large share of plans was financially dominated—i.e., more expensive, in some cases substantially so, than other plans regardless of eventual health spending or attitudes towards risk (e.g., selecting a $750-deductible plan over the otherwise same $1000-deductible plan would result in $528 higher premiums for a maximum cost-reduction of $250).

One of our central contributions is to document that a majority of employees selected dominated insurance. These choices were costly—dominated plan choice resulted in excess costs equivalent to 50 percent of the premium of the otherwise identical high-deductible plan—and regressive, because low-income employees were more likely to choose dominated plans. Moreover, the large majority of employees, and particularly those of low income did not switch into cheaper plans the following year, suggesting the consequences of poor choice amplify over time. To better understand employees' underlying motives, we administered an online experiment in which several thousand participants made hypothetical insurance decisions from a stylized representation of the real-life choice environment. We corroborated the validity of the vignette-centric experimental paradigm by documenting a striking correspondence between baseline choices in the lab and those of employees in the field. We then analyzed how participant choice

responded to experimental variation in the size, complexity, and clarity of the menu. The experiments indicate that dominated plan choice emerged not from menu complexity or preferences for avoiding high deductibles but through fundamental misunderstanding of cost-sharing.

Contributions. The central contribution of the paper is to leverage an unprecedented empirical setting to provide the first systematic evidence that the active choice of insurance violates predictions of expected utility, as well as some of the strongest evidence to date of regressive cost of poor choice.[1] For the economic literature on insurance, these findings suggest that the presumed likelihood of market unravelling due to adverse selection may be overstated and that such unraveling may depend on consumer sophistication as much as adverse selection. Moreover, the widespread violation of dominance points to a fundamental misspecification of the model used by economists to estimate risk preferences and to estimate welfare.

From the perspective of policy and consumer welfare, the pervasiveness, persistence, and regressivity of costly poor choice implies the potential welfare loss associated in settings where consumers face a menu with actuarially expensive high-coverage plans (regardless of dominance). The findings also challenge the rationale for increasing welfare via the expansion of choice in the context of large insurance marketplaces (ignoring potential cost-benefits resulting from increased plan competition). One could interpret these findings as emphasizing the need for decision aids and/or standardized exchanges that clarify potential cost distributions. However, I also see the results as emphasizing the benefits of strategically restricting choice options or encouraging the development of fundamentally simpler insurance plans that consumers actually understand (a point elaborated upon in **Bhargava** and Loewenstein 2016, AER P&P). Practically, the paper led to conversations with the CMS about the design and marketing of ACA exchanges, prompted a 2017 workshop at the University of Connecticut where representatives from the insurance industry convened to discuss the paper's implications for future product design, and compelled the firm providing us data to replace the build-paradigm with an HDHP/HSA menu for its 51,000+ benefit-eligible employees.

The paper won the Best Publication Award from the Behavioral Science & Policy Association in 2018 and was a finalist for the Annual Health Care Research Award presented by the NIHCM. We summarized the findings of this paper and the implications for policy in the *NEJM Catalyst* (Loewenstein and **Bhargava**, 2016). The paper was presented at the NBER meetings and at BEAM.

2.2.2 New Evidence on Motives for Financial Risk-Aversion

BLS 2017 illuminates how empirical investigation of financial risk-taking in insurance markets is likely confounded by consumer misunderstanding. While researchers have generated a number of insights into risk-taking in other field contexts, such contexts may be similarly challenged by limited decision-maker understanding (e.g., asset allocation, options trading), limited visibility into decision-maker beliefs (e.g., asset allocation, gambling) or limited generalizability to economically important contexts (e.g., high-stakes decisions in game shows). In a new paper, my former advisee, Timothy Hyde, and I sought to provide evidence on the prevalence of, and motives for, financial risk-taking from an empirical setting that uniquely avoids many of these empirical challenges (**Bhargava** and Hyde 2022, planned submission to *Review of Economic Studies* [BH 2022]).

The paper describes an analysis of the choices, productivity, and beliefs of 20,133 employees across 18 large North American firms participating in a simple, all-or-nothing, goal-rewards program, called GoalQuest (GQ).[2] Designed by a behavioral consulting firm seeking to encourage productivity, the program requires participating employees to self-select a productivity goal for the one- to three-month duration of the program from a menu of three personalized options. Critically, each goal corresponds to an often substantial all-or-nothing reward and while goals increase linearly, the value of rewards increase non-linearly. For example, an employee at a call center might select from the following menu of goals (Goal 1: 100 calls; Goal 2: 110 calls, Goal 3: 120 calls) and rewards (Goal 1: $100, Goal 2: $300, Goal 3: $600). If she selects Goal 2 and resolves 135 calls during the program, she only earns $300; if she selects Goal 2 and resolves 105 calls, she earns nothing. On an average, an employee could earn $466 (roughly six-percent of pro-rated annual salary) and our primary analysis involved $9.4 million in rewards. Due to the structure and simplicity of the program, its generalizability, and our ability to observe employee beliefs, the setting offers a rare

[1] At the time of publication, the other example of financially dominated plan choice was Handel (2013)'s formative analysis of employer-sponsored plans where such choice emerged through plan inertia and relative changes in plan prices.
[2] George Loewenstein generously facilitated contact with our partner firm (BIW) and access to this remarkable data.

opportunity to study decisions under risk and uncertainty. For example, with respect to generalizability, the sample features a diverse set of employees participating in the program at an estimated rate of 98 percent; these employees engage decisions whose value ($69 to $4,500) spans the range of many economic decisions of interest. Most importantly, our partnership with the firm led to the adoption of an "enhanced" enrollment protocol that asked employees to forecast their likelihood of attaining each goal (after goal choice), thereby permitting us to incorporate contemporaneous employee beliefs into our analyses.

The analyses generated three primary insights about financial risk-taking in the field. First, employee goal choices implied substantial risk aversion and choice diversity, patterns that persisted across economic stakes and employee tenure. About one-half of employees selected a lower goal than predicted by the expected-value maximizing benchmark assuming rational expectations (for most, Goal 3), resulting in an average unrealized gain equal to 30 to 58 percent of potential reward value. Second, conservative goal choice could not be explained through any set of plausible utility-based risk preferences. While assuming risk aversion at the upper bound of plausibility moderately reduced the share of employees for whom the benchmark deemed Goal 3 as optimal, it did not increase the share of explained goal choice. Finally, departures from the EU framework—biased beliefs, non-linear decision weights, and gain-loss utility—routinely discussed in the literature do not help to explain choice in this setting. For example, while a systematic bias in beliefs privileging lower goals could, in theory, produce conservative goal choice among otherwise utility-maximizing employees, employees were appreciably *overconfident* in both their relative and absolute beliefs of high-goal attainment. As another example, while the literature has suggested that gain-loss utility might help to explain risk-averse choices, tests of 70 models across an exhaustive range of loss aversion parameters, functional specifications, and candidate reference points, indicate no systematic improvement in explanatory power. We also replicated the same pattern of risk-aversion from the field in an incentive-compatible online goal-reward paradigm with dollar-denominated rewards, verified comprehension, and minimal scope for signaling, reputational concern, or effort costs.

Heuristic Explanation for Risky Choice. Prompted by pilot studies exploring choice processes, we proposed a novel explanation for risk averse (and heterogeneous) goal choice. The proposed "pairwise heuristic" stipulates employees select a goal through a succession of approximate comparisons between proximal pairs of goals. Critically, we assert that the contingent inference required by the pairwise comparisons is subject to systematic bias, leading employees to underestimate the relative likelihood of attaining the riskier goal, thereby increasing their chance of choosing the lower, less-risky, goal.[3] We tested the proposed heuristic with a new experiment in which we asked participants to make hypothetical goal-reward choices from representative menus. The experiment documented participant reliance on proximal pairwise comparisons and a systematic, and often substantial, bias in estimates of contingent likelihoods across domains like weather, sports, goal attainment.[4] A simple additively-linear choice model predicts that the absence of this bias would increase expected-value maximizing choice by 37 percent. The experiment also documented whether the optimality of participant choice increased when made from a "de-biased" menu that displayed accurate contingent likelihoods associated with each goal (e.g., the likelihood of attaining goal 3 conditioned on attaining goal 2). Relative to a baseline menu displaying average non-contingent likelihoods of attainment, choice from the de-biased menu resulted in a 48 percent increase in optimal choice share. As further evidence for the proposed heuristic, choice from the menu displaying non-contingent likelihoods looked nearly-identical to choice from a menu displaying no likelihoods and to a menu displaying contingent likelihoods parametrically adjusted for the hypothesized bias. Finally, in assessing descriptive accuracy, we found that the heuristic explained 54 to 60 percent more choice than the subjective EV benchmark in the lab and, using a parametric representation of the heuristic, 26 to 46 percent more choice than the same benchmark in the field.

Contributions and Follow-up Research. This paper contributes to the literature on financial risk-taking in the field by documenting substantial risk-aversion in a simple setting with wide-ranging economic stakes and high-compliance. Because of rare access to data on decision-maker beliefs, it also document the limited explanatory power of traditional, and behaviorally-modified, utility-based explanations. In addition, the paper compellingly shows that a non-

[3] The heuristic implies an employee deliberating between goals 2 and 3 would roughly assess whether the expected potential gain from the higher goal (the difference in reward weighted by the perceived conditional likelihood of attainment, $\Delta r_{3,2} * \hat{s}_{3|2}$) exceeds the potential loss from not attaining the higher goal ($\hat{s}_{\neg 3|2} * r_2$). The employee underestimates the likelihood due to insufficient adjustment for the contingent nature of the comparison.

[4] As an example of the latter, relative to the probability implied by non-contingent elicitations, participants substantially underestimated the likelihood that Steph Curry would score 30+ in his next NBA game if they knew, with credible certainty, Steph would score at least 20+ points.

trivial, and possibly far more substantial, share of employees relied on a heuristic choice-strategy involving a bias in contingent inference. A practical question is whether the proposed heuristic could help to explain risk aversion and the unexplained heterogeneity in risky choice elsewhere? In a final experiment in the paper, we show that demand for home insurance, from menus informed by field research (Sydnor 2010), conforms far more strongly to EU-benchmarks (both in implied risk-taking and heterogeneity) when made from menus discouraging contingent comparisons, suggesting that seemingly anomalous insurance choice may reflect heuristics with contingency neglect. Finally, the proposed heuristic, which has not been discussed in the literature, adds to a short list of heuristics advanced as explanations for risk-averse behavior (e.g., the priority heuristic, Brandstatter et al. 2006). Moreover, the sizeable bias in pairwise contingent belief at the heart of the heuristic, that we demonstrate across domains and elicitations, has not been previously documented. In future work, I hope to clarify the processes underlying the bias, to derive a more sophisticated functional representation than used in the paper, and to better understand how this phenomenon relates to existing models of biased inference and probability neglect (see Benjamin 2019; Sunstein 2003; Fox et al. 2015).

2.2.3 Policy and Program Design

As noted, a policy implication of the preceding work is to emphasize the importance of how insurance exchange design. Given the roughly contemporaneous launch of the insurance exchanges created by the ACA, in two papers, my colleagues and I attempted to estimate the financial consequences, and likelihood, of sub-optimal consumer choice by ACA enrollees (**Bhargava** and Loewenstein, *JAMA* 2015; **Bhargava**, Loewenstein, Benartzi, *BSP* 2017). First, we used data on the 39,885 plans offered in the inaugural year of the 34 federally-sponsored exchanges to project the total health spending for representative consumers under plans representative of each of the four cost-sharing tiers within which plans on the exchange were organized. The analysis suggested that selecting the "second-best" plan tier, relative to expected need, would lead to excess spending equivalent to 13 to 37 percent of pre-subsidy premiums. We then assessed how likely ACA enrollees were to make inefficient choices from the exchanges, which by legislative decree, denoted actuarial tiers with metal-labels (Bronze, Silver, Gold, Platinum). Through hypothetical choice experiments from menus reflecting actual ACA plans and pricing, we found that two-thirds of subjects chose plans incommensurate with their self-reported health needs and risk attitudes, resulted in excess spending equivalent to 24 percent of average premiums. We additionally found that the metal labels were no more effective at promoting efficient choice than generic labels and substantially less effective than labels emphasizing anticipated health care need. Once again, these papers highlight a recurrent theme regarding the critical importance of understanding context-specific decision-processes for optimal program/policy design.

2.2.4 Work-in-Progress

I briefly describe three ongoing projects that build on the preceding research (all with anticipated completion dates in 2023). A first paper, Heuristics Gone Awry, revisits the firm setting examined in BLS 2017 to explicitly identify the (heuristic) choice strategies responsible for plan decisions (**Bhargava**, Loewenstein, and Sydnor, anticipated 2023). While BLS 2017 documents insurance choice that cannot be explained using existing economic frameworks, as evidenced by systematic "lumpiness" in choice patterns in the lab and field, employees did not choose plans at random. BLS 2023 explores the choice strategies employees adopted to select plans. It leverages a rather remarkable natural experiment—prior to abandoning the build-your-own-plan paradigm, our partner firm actually *expanded* its menu from 48 to 60 plans, retaining its dominated nature—and a more expansive set of hypothetical choice experiments to test candidate heuristics (including the pairwise heuristic from BH 2022).

A second paper, Behavioral Erosion of Private Information, explores how behavioral biases can systematically erode private knowledge of risk and how such erosion could help to explain well-documented empirical puzzles in the insurance literature (**Bhargava** and Hyde, anticipated January 2023). The paper exploits access to an unprecedented combination of telematics data detailing the actual driver risk for several thousand enrollees of a user-based insurance program administered by a large US auto-insurer linked to detailed survey-data (for the same enrollees) capturing forecasts of risk, measures of risk preference, hypothetical insurance choice, and several decision-making assessments. The analysis reveals that while drivers have private knowledge of their own relative risk—i.e., risk that cannot be predicted from legally-permissible insurer rating variables, most drivers nevertheless underestimate their insurance risk. Using driver forecasts and surveyed assessments of biases, we estimate a "behaviorally-adjusted" risk forecast to show

bias erodes average and marginal private knowledge of risk (two constructs defined in the paper) by 31 to 53 percent. Finally, we use hypothetical insurance choices from the same sample to show how behavioral erosion offers a possible resolution to empirical puzzles of adverse selection (Cutler, Finkelstein, McGarry 2008).

Finally, in a third paper, Gender Differences in Risk-Aversion in the Field, I plan to follow-up on an intriguing pattern in the data analyzed in BH 2022—a substantial gender difference in financial risk-taking, whereby women select significantly more conservative goals (**Bhargava** and Hyde, anticipated Summer 2023). This difference cannot be explained through differences in ability, utility-based risk preferences, or differences in the accuracy of beliefs (men and women are substantially, and similarly, overconfident). I tentatively hypothesize that the gender difference actually reflects the differential adoption of heuristic choice strategies.

2.3 Retirement Saving in the U.S.

The standard life-cycle model of saving struggles to explain several empirical features of how US employees save. For example, many employees undersave for retirement despite access to tax-advantaged 401(k) plans with matching incentives, are not highly responsive to changes in the generosity of the match (Madrian 2013), and routinely express the intent to save more but fail to follow through (Choi et al. 2002). Equally puzzling for standard theory, employees do respond to non-economic features of plan structure such as the presence of automatic enrollment (AE) (Madrian and Shea 2001). It bears noting that the risk of retirement insecurity (a circumstance where a retiree is unable to cover basic expenses without resort to means-tested public assistance, a return to work, or bankruptcy) likely extends to a higher share of actual 401(k) *enrollees* than believed. As motivating analysis in a new paper, my colleagues and I attempted to characterize the state of retirement preparedness across 401(k) plan enrollees (**Bhargava**, Conell-Price, Mason, Benartzi, 2021, Save(d) by Design, BCMB 2021). The evidence offers a rare snapshot of contemporaneous preparedness informed from simulations of potential asset accumulation under widely-varying assumptions, informed by administrative records from 186,000 enrollees across 840 AE plans. Our preferred estimates imply 44 percent of current enrollees face a non-trivial risk (arbitrarily defined as 25+ percent) of insecurity, a likely conservative estimate given the pre-pandemic exercise ignores savings leakage due to job loss, loans, etc.

In this section I describe three papers that adjudicate leading demand-side explanations for undersaving and offers a potentially unifying explanation for demand-side puzzles; identify a supply-side dynamic that, if addressed, could improve retirement preparedness to a degree comparable to the introduction of automatic enrollment; and advance a portfolio of near- and long-term strategies for policy reform and increasing engagement. Collectively, these papers describe findings from field experiments across 899 firms, hypothetical enrollment decisions by an additional several-thousand employees, and surveys of dozens of high-ranking policymakers and hundreds of plan executives.

2.3.1 Demand-Side Barriers to Saving ("Serenity Now")

Economists have advanced a number of potential departures from the standard economic framework—or psychological frictions—to explain employee undersaving (and other empirical saving puzzles). Four of these have come to occupy a central role in the academic literature, policy discourse, and industry best-practices: *Retirement Literacy* (low financial literacy or biases in retirement-relevant beliefs leading an employee to underestimate the need to save), *Plan Confusion* (underestimation of plan eligibility or match generosity may lead an employee to underestimate value of enrollment), *Enrollment Complexity* (perception of administrative enrollment as involving prohibitively high psychological and/or time-costs), and *Present Focus* (low-saving due to the tendency to disproportionately privilege immediate, relative to delayed, flows of experienced utility).

Despite the regularity with which these four frictions are discussed, evidence as to their causal role in the employee saving remains scarce (Beshears et al. 2018). In a paper with a former advisee, Lynn Conell-Price, we attempted to provide such evidence via a high-compliance field experiment targeting 1,137 undersaving, 401(k) plan-eligible, employees at a large US firm with a generous plan match (**Bhargava** and Conell-Price, 2022; Serenity Now; under review *American Economic Review*, BC 2022). We administered the field study by inviting qualified employees to participate in an online survey marketed as an opportunity to provide confidential workplace feedback. The first module of the survey diagnosed the employee-specific incidence of each candidate friction while a second module promised employees a personalized assessment of their retirement preparedness. The assessment truthfully conveyed the employee was not "on track" for retirement security, advised the employee to increase their contribution rate,

provided guided instructions to do so, and, finally, asked about future intentions to save. To test each friction, the module also experimentally varied the presence of (1) A personalized saving recommendation, (2) Information clarifying the magnitude of the match, and (3) A small, but immediate, reward ($10 Amazon gift card) to encourage employees to visit the enrollment portal and consider the decision to save more. Embedding the field experiment within a survey of beliefs and decision-making permitted us to estimate the baseline prevalence of each friction (and its naïve correlation with saving), the average effect of reducing a specific friction on saving, and the differential effect of reducing a specific friction on saving across baseline incidence.

We report four primary findings from the field experiment: (1) While nearly one-half of employees underestimated how much they should save such deficits *do not* appear to cause undersaving. That is, providing a concrete, personalized recommendation (that verifiably improved belief accuracy) had a precise, small, effect on saving, even among employees with low baseline literacy. (2) Employee confusion may explain a significant share of undersaving—not confusion about eligibility (almost all employees knew they were eligible) or match generosity (employees did underestimate the match, but clarifying its size did not increase saving). Instead, in an unplanned analysis, we found that 37 percent of non-participants mistakenly believed themselves to be enrolled. Presumably confused employees assigned to the small reward (and thus more likely to observe their actual status), were three times as likely as counterparts to increase saving. (3) Little evidence suggests perceived enrollment complexity deterred saving in this setting, a result that we speculate reflects the ease with which employees can enroll, or adjust contributions, at digitally administered plans. (4) As evidence for the importance of present focus, 8 to 16 percent of employees increased their contribution, often substantially, in response to the $10 reward, despite not responding to clarification of the larger, but delayed, match. Moreover, employees tagged as present-focused in the survey were 2 to 3.5 times more responsive to the reward than counterparts. For those who hadn't exhausted the match, the response to the $10 implied an average gain from the match of $2,632 through the next year. Present focus was also indicated by the larger share of employees who, at baseline, failed to take-up the match despite a stated preference to save, confirmed knowledge of the match, and the absence of illiquidity.

Explaining Employee Present Focus. What drives the present focus of employees? Given its precedence in the literature, we first considered a beta-delta model of present bias, adapted from DellaVigna (2018). We show through calibrations that such a model cannot plausibly rationalize enrollment delay for more than a few days—even allowing for significant psychological costs of enrollment. While an established approach in the literature is to explain lengthier delays by assuming employee naïveté, most employees in our field experiment reported an intent to save more in the future, but only after a delay of weeks to months. Even without data on future intentions, the calibrations suggest the improbability of explaining saving in response to the $10 reward but not match clarification. We conclude by proposing a novel belief-based hedonic account of present focus informed by an intriguing empirical correspondence involving saving and financial anxiety, an exploratory friction assessed in the survey: Most employees reported substantial anxiety about their financial situation yet were optimistic about their anxiety subsiding in the intermediate (weeks to months) but not immediate (days) future. This parallels the dynamic of low present saving and delayed intent to save more in the future. The proposed model ("Serenity Model") suggests that anxiety felt by undersaving employees generates a hedonic cost to engaging consequential financial decisions, such as plan enrollment. Critically, employees believe their anxiety to be temporary and have well-defined beliefs about their eventual transition to a low-anxiety state. The model predicts how an otherwise rational employee with anxiety could delay enrollment if the expected benefit of delay (less anxious enrollment) exceeds its cost (the foregone plan match). The framework also offers a lens to understand how a micro-incentive—which we define as a small, immediate, unexpected, incentive encoded and processed through the dopamine reward pathways of the brain—could motivate employees not through its monetary value but by reframing enrollment as an exercise in reward-seeking (an interpretation for which we cite neuroscientific support and present experimental evidence). Calibrations show how the model can explain both lengthy enrollment delays and future saving intent in the paper without implausibly high costs of enrollment.

Policy Implications. The paper offers several practical lessons for increasing engagement/saving of financially at-risk employees. In the near-term, the evidence challenges the presumptive usefulness of strategies relying on personalized guidance or financial education at the time of enrollment or campaigns aimed at clarifying, or even increasing, the match. Instead, the evidence points to the potential efficacy of participatory micro-incentives and, suggestively, reforms that increase awareness of plan status. A survey of dozens of high-ranking (federal/state)

government and industry stakeholders in 2021 confirmed that our findings signal a departure from consensus priorities for reform. In the long-term, if our hedonic account of present focus does describe the behavior of a non-trivial share of employees, it suggests the need for more fundamental reform that better aligns plan structure with the underlying psychology governing the decision to save. An example of such reform are dual-account proposals that would direct initial contributions to a liquid buffer account before automatically transferring above-threshold savings into a less-liquid account. Such plans have been recently advocated by concerned over short-term illiquidity (e.g., Beshears et al. 2015; 2020). Our findings offer a psychological rationale for such proposals in that, for many employees, addressing near-term financial anxiety may be a precursor for long-term saving.

Contributions and Follow-up Research. I consider this paper, along with BM 2015 and BLS 2017, as my most important research to date. Methodologically, it is the only field experiment of which I am aware on retirement savings to embed person-specific measures of multiple frictions within experimental tests of those same frictions. And to avoid challenges of drawing population-level inferences from the low- and potentially unrepresentative response often associated with many disclosure experiments, we engineered this study for high compliance (most employees exposed to a treatment likely engaged the treatment, as indicated by minimal dropout and built-in attention checks and belief elicitations). Practically, the findings allude to two novel strategies for increasing engagement among a population often seen as unresponsive to traditional incentives (micro-incentives and efforts to combat enrollment status confusion) and offers a narrow but forceful challenge to the popular practice of delivering financial literacy/education at the time of enrollment as a means of increasing engagement.[5,6] It also offers a micro-foundational rationale for more structural reform involving dual accounts. Theoretically, the Serenity Model offers a hedonic, belief-based, alternative to preference-based explanations for present focus I hope to explicitly test and develop in future research. In the paper, we describe how the model could be used to explain several savings puzzles (undersaving, match insensitivity, the efficacy of AE, persistent gaps in actual and intended saving). It also provides a framework for understanding the outsized influence of small participatory incentives. Finally, the phenomenon of high present anxiety and delayed expectation of hedonic relief that serves as the basis for the model has not, to my knowledge, been discussed in the research on the optimism bias or hedonic forecasting. I hope to summarize this effect (since replicated across a number of contexts), along with proposed mechanisms, in a paper aimed at the psychology/JDM literature. The paper, which is under review, was presented at BEAM 2020 and has elicited encouraging reactions from those I greatly admire in the field.[7]

2.3.2 Supply-Side Barriers to Saving ("Save(d) by Design")

While the prior paper investigated the demand-side barriers to saving, a contemporaneous project identifies a feature of the behavioral design of 401(k) enrollment that, if addressed, could lead to an increase in employee preparedness for retirement comparable to the introduction of automatic enrollment (**Bhargava**, Conell-Price, Mason, Benartzi, Save(d) by Design, planned December submission to *American Economic Review*). We begin the paper with two motivating considerations. First, we present simulation-based projections of retirement risk, described earlier, indicating that the risk of retirement insecurity extends not just to those without 401(k) plan access or those who have not enrolled in the plan to a large share of actual 401(k) enrollees (the same analysis also documents unexplained differences in average preparedness across observably similar employees at economically similar plans). Second, we posit that one factor that may help explain the pattern of high (and variable) enrollee retirement risk in AE plans are the non-economic features of plan design that substantially shape an employee's initial enrollment decision. Practically, newly-eligible employees in an AE 401(k) plan either enroll through inaction or by actively visiting an online enrollment interface to select one of three options: confirm automatic enrollment at the default rate, personalize enrollment at an adjusted rate, or decline enrollment. Employees deciding to personalize enrollment then proceed through a short webflow to adjust their contribution.

---

[5] While establishing the breadth of enrollment confusion is for future research, we found comparable degrees of self-reported uncertainty about 401(k) status in a separate survey of 500 employees. We speculate such confusion may arise from the broader complexity of the benefit program landscape.
[6] Of course, our findings do not rule out the possibility that such interventions may increase saving through other channels such as persuasion, social pressure, etc., nor do they rule out potential benefits of financial education outside of engagement
[7] In an unsolicited email, Cass Sunstein remarked "just finished your serenity now paper. Wow. It is amazingly good! Learning so much from it." David Laibson remarked that it was "brilliant" and full of insights. And Larry Katz referred to it as "top-notch" and "first-rate" work.

There are reasons to suspect an employee's active decision to personalize enrollment may have an outsized influence on their long-run financial security. This is because default rates are set far below recommended targets and below most plan match thresholds and employees do not often adjust their contributions year-to-year, so small increases in initial contribution, through personalized enrollment, can lead to large increases in eventual savings. As illustration in cross-sectional analyses of our data, personalized enrollees contributed at a rate *more than double* the default, leading to a degree of projected preparedness far exceeding their automatically enrolled counterparts (one, of course, presumes this reflects employee selection). Notably plans, within and across provider, vary considerably in what we refer to as the *psychological design* of (initial) enrollment—that is, plans vary in the intensity and frequency of marketing efforts and reminders to encourage newly-eligible employees to actively enroll and they vary in the visual appearance, usability, and interactivity of the online enrollment interface.

Research Overview. Given the presumptive importance of initial enrollment for retirement security, our aim in this paper was to test the influence of psychological design on the initial enrollment of newly 401(k) plan-eligible employees and to clarify the implications of these findings for policy, welfare, consumer protection, and decision theory. Our central evidence draws from three field experiments a large 401(k) record-keeper, that randomized several aspects of the psychological design of the otherwise standardized online interface from which 8,565 employees from 500 AE plans made their initial enrollment decisions. Specifically, we varied non-economic aspects of design that plans and providers routinely vary (the use of "traffic-light" colors, declarative versus descriptive headlines, standardized descriptive language, spotlighting previously communicated plan detail like the default rate, etc.). The studies were administered to the near-universe of small-to-midsize market plans for which our partner, Voya Financial, was the record-keeper from 2016 to 2018. To investigate the mechanisms underlying response to the treatments, we supplemented the field studies with evidence from 6,871 employees asked to make hypothetical enrollment decisions from experimentally varying online interfaces programmed to look and function like their real-life analogues. Finally, to better understand the sophistication of plan sponsors, we surveyed several hundred plan administrators to test their ability to forecast the influence of design and to identify successful design elements.

Our analyses produced four primary findings. First, we document how modest changes to the psychological design of the interface (*enhanced design*) resulted in sizable increases in personalized enrollment (+15% from a 0.60 baseline), full match take-up (+16%), and average savings (+13%) relative to the commercial baseline (*basic design*)—effects equivalent to that predicted from a 68 to 74 percent increase in the modal match. The influence of enhanced design on increased saving was largely invariant to the financial stakes of the decision (as implied by the difference between the match and default rate). Second, we show that, remarkably, marginal personalized enrollees appear to more than double their contributions relative to the default, just like their inframarginal counterparts. As a result, accounting for the plan match and inertia, we project that that the shift from automatic to personalized enrollment, due to design, would yield several additional years of retirement security, particularly for employees who begin to save at an older age. Third, lab evidence indicates that design did not affect enrollment by shifting preferences or decision-relevant beliefs (or through other behavioral frictions such as inattention, confusion, or distrust). Instead, it seems, for many employees, enrollment emerged from a heuristic process in which design shifted affective appraisals (that is, the influence of design was mediated by shifts in anxiety associated with personalized enrollment). Finally, our industry surveys suggest most plan administrators underestimated the potency of design and couldn't identify optimal design elements.

Contributions and Implications for Policy/Welfare/Theory. This paper describes how the behavioral structure of 401(k) enrollment may profoundly shape retirement preparedness—to a degree that compares with the introduction of automatic enrollment. I describe five specific contributions pertaining to policy/welfare/theory. First, our motivating estimates of retirement preparedness from an analysis of administrative records offers the most direct evidence that the risk of retirement insecurity likely extends to millions of actual 401(k) enrollees—a risk underestimated by leading stakeholders (as evidenced through the same stakeholder survey described in BC 2022). Our use of simulations, annuity-based mapping of savings to retirement income flows, and insecurity spending thresholds reflects a novel strategy to characterize retirement risk, an arguably more welfare-relevant outcome than the average ratios of income replacement often discussed in the literature. Second, the large effect of design on saving is less important as a demonstration of efficacy of a particular nudge (excellent prior work has shown how simplification, reminders, and framing can affect saving in the field), but rather as a demonstration of the potentially first-order influence of (digital) design on initial enrollment outcomes and retirement preparedness. The potency of design stands in contrast to more

expensive engagement strategies such as increases to the plan match or plan marketing. A third contribution is to document the marginal-inframarginal equivalence in saving among personalized enrollees. Our research setting uniquely permits us to infer marginal behavior under reasonable assumptions since, unlike most settings, we observe an extensive-margin decision to personalize enrollment followed by an intensive-margin decision of how much to contribute. An implication of this equivalence is that consumers do not have well-formed and stable preferences for something as consequential as saving—a possibility with further implications for economic models of enrollment and welfare analyzes of saving interventions. A fourth contribution is to highlight the implications of psychological design for consumer protection and, perhaps, definitions of fiduciary responsibility. Given the potency of design, the lack of sophistication among plans as to such potency, and the substantial commercial variation in design across providers and plans—variation far more pronounced than the modest differences we tested in the paper—suggests that design could result in large and unintended variation in preparedness. In the paper, we present experimental evidence alluding to large differences in enrollment outcomes associated with the varying enrollment designs used by providers/plans.

Finally, I see this research as an early step in examining the potentially profound influence of digital design on an array of financial and consumer decisions from digital settings. At first glance, design could be seen as a type of (digital) nudge—small, choice-preserving, potentially welfare-enhancing, modifications to a decision environment. Practically, economic research on nudges has documented the importance of defaults, complexity, numerical framing, reminders, and salience across a range of economic decisions and has interpreted nudges as influencing choice by addressing cognitive limits pertaining to self-control, attention, computation, and understanding (e.g., Datta and Mullainathan 2014). However, the factors driving behavioral response to design in the present setting seem more akin to the hedonic and motivational channels commonly invoked in research on reward-seeking and on gamification. If one interprets design through this lens, then the implications for design for policy and consumer welfare should differ from choice-preserving, and largely predictable, nudges. I hope to continue to explore the theoretical and practical importance of digital design, and the related topic of gamification, in future research.

2.5 Other Topics

2.5.1 Emotion and Well-Being

One focus of my research is investigating the causes of well-being, documenting the descriptive experience of emotion, and using such descriptions to better understand emotion theory, and exploring how the experienced-decision utility gap can inform economic analyses. I briefly describe four papers in this section and note that I have a number of future analyses planned. At the center of this agenda is an extraordinary proprietary dataset, collected by a now-defunct marketing firm, of unprecedented scope and detail. The data describes the in-the-moment mood, arousal, and 17 specific emotions (e.g., hope, fear, anger, boredom), details of social, locational, and activity time-use reported on mobile devices every 30-minutes over a 10-day period. The data additionally captures TV, online, and media consumption and details purchasing activity. The sample spans a diverse set of 7,122 US adults across 2.2 million observations for whom I observe dozens of person-specific demographic, psychographic, and lifestyle descriptors. Not only do in-the-moment measures of emotion correlate with physiological hedonic markers, they are less sensitive to biases (imperfect recall, scale adaptation, demand effects) than the retrospective/evaluative measures typically used by researchers.

A first pair of papers engages a longstanding empirical puzzle in the well-being literature: Children do not appear to increase parental well-being. This consensus account, reflected across scores of papers varying in data and methodology, is difficult to reconcile with evolutionary directives and lay-beliefs. A well-publicized rejection of the consensus (Nelson et al. 2013) relied on an analysis heavily confounded by factors including marital status—a point that I made in a reanalysis of their data **(Bhargava***, Kassam, and Loewenstein, *Psychological Science* 2014). In a new paper, I weigh in on this debate with novel evidence from the aforementioned dataset **(Bhargava**, The Hedonic and Time-Use Consequences of Parenthood, planned submission to *PNAS*). My analysis reveals a moderate, positive, and statistically significant, overall effect of children on parental mood and emotion—entirely driven by a large positive effect of children on mothers (approaching the estimated positive effect of marriage, the most robust source of well-being in the literature). Decomposition analyses indicate that one can attribute 30 to 42 percent of the observed hedonic gender gap to systematic differences in how mothers and fathers allocate their time and men and women appear to enjoy

time with children in equal measure. Intriguingly, among those with social and/or financial resources, children increase parental well-being for both men and women. I speculate as to how to reconcile these data with the existing literature.

In a second paper (**Bhargava*** and Kassam, Love and Gender: An Empirical Account, under review at *Psychological Science*), I present evidence on the prevalence, correlates, dynamics, and hedonic consequences of experienced love. The paper supplements the aforementioned data with a second large experience sample. The paper describes analyses alluding to the similarity of experience across gender and to its functional nature—i.e., men and women report comparable degrees of (passionate) partner love, such love declines for both genders after early marriage but rises after prolonged partner separations, and greater reported love predicts substantially higher in-the-moment well-being across gender. The gender differences we do find—women report child love more frequently than men and men exhibit a less severe drop in partner love (and possibly passion) over marriage (a pattern suggestively mediated by children)—also corroborate a functional account of love that recognizes the varying role of men and women in relationships. Finally, a third paper, led by a graduate student in our department (Chin, Markey, **Bhargava**, Kassam, Loewenstein, *Emotion*, 2017) offers a descriptive account of when and where people experience boredom and attempted to interpret this account with respect to competing theories as to the causes of boredom.

2.4.2 Translating Effects from the Lab to the Field

Finally, I describe two projects that demonstrate a long-standing interest in understanding how well-established experimental phenomena translate to complicated field settings with incentivized outcomes, repeated exposure and the potential for learning, representative populations, and the opportunity for people to engage in substitutive and compensatory actions. A first paper, Driving Under the (Cellular) Influence, exploited a natural experiment and nearly two dozen datasets to generate causal evidence on the link between driver cell phone use and crash risk (**Bhargava** and Pathania, *American Economic Journal: Policy* 2013). At the time of our research (a period predating smartphones), the overwhelming consensus of 120+ studies, including many lab studies involving driving simulators, was to equate the risks of cellular use while driving to drunk driving, leading numerous states to institute partial or full bans. We sought new evidence on the question by first documenting the exogenous rise in phone use by drivers, prompted by the switch from peak to off-peak minutes at 9PM on weekdays that characterized most cellular plans at the time. We then analyzed data on millions of crashes to see whether the rise in cellular use led to a corresponding rise in crashes. Our estimates indicated no increase in crashes and were sufficiently precise to reject consensus estimates from the literature. Beyond leveraging an attractive natural experiment, the paper is notable in relying on over twenty datasets summarizing millions of crashes, pricing plans for tens of thousands of users, and hundreds-of-thousands of cellular calls. The data on calls included perhaps the first academic appearance of data on cellular signals and their routing patterns across towers, permitting us to infer whether a signal emanated from a moving vehicle (data reflecting over a year of approaching every major cellular provider in the country). Beyond challenging the rationale for cellular bans, the paper highlights the perils of naïve translation of even well-established results from the lab to the field. We note that such translation could ignore factors such as compensatory behavior by drivers (e.g., drivers compensate for the dangers of cellular distraction by driving more carefully). Our contrarian findings were, in my view, corroborated by a series of RCTs administered by the NHTSA a few years later. Prior to publication in the *AEJ: Policy*, the paper generated 9 referee reports across two submissions to the *QJE*, a negative decision the editor characterized as coming down to a "toss-up."

A second project in this category sought to understand whether sequential decisions in the field reflected the cognitive and perceptual biases implied by decades of experimental evidence. Specifically, it has been well-documented that sequential judgment and choice in the lab is often biased by perceptual contrast effects (a bias whereby evaluation of contemporaneous stimuli is negatively contrasted with preceding stimuli). However, the experimental literature is less clear as to when such contrasts effects emerge (as opposed to, for example, assimilation effects), the magnitude of bias, the evolution of any bias over decision experience, etc. In work with Ray Fisman, (**Bhargava** and Fisman, *The Review of Economics and Statistics*, 2013), we sought to present the first empirical evidence for sequential contrast effects in a high-frequency decision-setting. Specifically, we advanced a conceptual framework from which to cleanly distinguish perceptual contrast effects from other mechanisms such as learning or quotas. We then analyzed data from a quasi-random sequential choice setting, speed-dating, where we had access to beliefs and decisions of several thousand decision-makers. The analysis revealed significant, short-lived, contrast effects in partner assessments (an effect largely driven by men). Such bias, not undone by decision experience, significantly distorted dating decisions. This initial

evidence for contrast effects in a high-frequency field setting presaged the possibility of similar bias in other settings such as hiring, medical diagnoses and investing (Hartzmark and Shue 2016). In unpublished work, I found evidence for contrast effects in the sentencing decisions of lower-court PA judges and in the evaluation of exams by teachers.

3. ADVISING, TEACHING, AND SERVICE

Apart from research, the aspect of my role that I enjoy the most is the opportunity to advise students and mentor junior faculty. My conviction as to the importance of compassionate and constructive advising grew from the generosity and career-shaping influence of my own cherished advisors, mentors, and senior colleagues. As a result, I have invested considerable time on these fronts, as evidenced by service on ten PhD student committee including chairing committees of four students and serving on three faculty reappointment committees. I have helped several PhD students prepare for the academic job market and developed a "job market boot-camp" for market-bound PhDs. My advising extends to undergraduates. I have worked with roughly 20-25 undergraduates through thesis advising, independent studies, or research assistantships and, in response to a perceived lack of career resources, I created a job market session focused on opportunities in consulting, policy, and behavioral science. I have coached several students for consulting interviews and have placed numerous students with interviews through industry/policy contacts.

While my philosophy as a teacher is to engage students through applied examples from policy and industry and my own rather impressive knowledge of Reality TV, and through one-on-one mentorship, my most meaningful pedagogical contributions likely come in the form of course development. I have developed three undergraduate courses during my time at CMU. A first, *Behavioral Economics in the Wild*, organizes relevant evidence from Behavioral Economics, psychology, and JDM, into three conceptual departures from the standard model (following Rabin, 1998; DellaVigna 2009)—non-standard preferences, non-standard beliefs, and non-standard decision-making. The course then works through the implications of these departures for policy/program design. A second, *Causal Inference in the Field* is an applied econometrics course distinguished by the number of in-class labs, applied group projects, and empirical case-studies using datasets intended to appeal to undergraduates. A third, developed with colleague Silvia Saccardo, is a unique project-based Capstone course in which seniors majoring in BE provide behaviorally-informed, evidence-based, counsel to professional clients on issues of strategic importance (our first client was the NHL's Pittsburgh Penguins).

Finally, I have striven to be a conscientious member of the department and the broader college. I have actively participated in several faculty hiring cycles within and outside my department, PhD admissions in all but one year when we did not admit a class, and was the founding Faculty Director of our recently launched BE undergraduate degree program. I have contributed to several more strategic efforts aimed at reforming departmental policies and practices, including leading efforts to renegotiate benefits for junior faculty and graduate students, helping to develop, market, and teach our department's Executive Education program in BE, and advocating for the creation of a system of "portfolio" advisors to better prepare graduate students for specific job markets. I have a courtesy appointment at Heinz College and serve as an affiliate for the Future of Work Initiative at the Block Center.