

Research and Teaching Statement

Daniel L. Chen

January 2023

1 Introduction

Law and economics is divided between the consequentialist view that optimal policy should be based on calculations of costs and benefits and a non-consequentialist view that policy should be determined deontologically: from duties we derive what is the correct law—what is right and just.

Are there deontological motivations, and if there are, how might one formally model these motivations? What are the implications of things like deontological motivations for economics methods and policy, and what puzzles can be explained with deontological motivations that one cannot with standard models? And how do deontological motivations interact with consequentialist ones in the production of justice? What is the impact of law and economics on justice? Do these motives provide a framework for improving the rule of law?

To answer these questions, my research has

1. **curated 12 terabytes of archival and administrative data on judges and courts** where normative commitments incubate; the data bridge machine learning, causal inference, and normative theories of justice regarding equal treatment before the law and equality based on recognition of difference
2. **developed a programming language to study normative commitments in experiments**, now used in over 23 countries, 10 academic disciplines, private and public sectors, and high schools
3. **spearheaded randomized impact evaluations to improve justice** with high-frequency administrative data in 17 countries

Some current themes on consequences, formation, and measurement of normative commitments (and applications in law) include:

- **Law and Development**—tracing the incentives that led to human rights violations
- **Markets and Morality**—how market forces interact with normative commitments
- **Behavioral Judging**—social and psychological, economic and political influences on legal ideas and production of justice
- **Law and Legitimacy**—role of legitimacy in legal compliance
- **Demography of Ideas**—economics of interpretation (hermemetrics) as a source of normative commitments
- **AI and Rule of Law**—leveraging normative commitments and AI to facilitate access to justice

My research has been accepted in leading economics journals (*American Economic Review*, *Econometrica*, *Quarterly Journal of Economics*, and *Journal of Political Economy*), leading computer science journals (*Journal of Machine Learning Research*), science journals (*Proceedings of the National Academy of Sciences*, *Science Advances*,

Nature Human Behavior), double-blind peer-review law outlets (Stanford-Yale Junior Faculty Forum and Law and STEM Junior Faculty Forum), 5 NeurIPS selections (Machine Learning and Law, Interpretable Machine Learning, CausalML, ML for Economic Policy, and AI for Credible Elections), and press outlets (*Washington Post*, *New York Times*, *Wall Street Journal*, and *Times of India*) and has been referenced in 2 National Academy of Sciences Study Reports (Deterrence and the Death Penalty (2012) and Decarcerating Correctional Facilities during COVID-19 (2020)).

The research has anchored successful grant applications with € 10 350 000 in grant budget awarded for “Origins and Effects of Normative Commitments”, “Positive Foundations of Normative Commitments”, “Digital Humanities: Legal Analysis in a Big Data World”, “The Impact of Justice Innovations on Poverty, Growth, and Development”, “Data AI and IE”, “Green and Digital Development”, “High-Dimensional Econometrics Applications in Law and Economics”, “Markets and Morality: Do Free Markets Corrode Moral Values?”, and "oTree: An Open-Source Platform for Online, Lab, and Field Experiments", and received support from The Alfred P. Sloan Foundation, European Research Council, Swiss National Science Foundation, Russell Sage Foundation, DFID, Google Inclusion, and Agence Nationale de la Recherche. My work has also been supported by the International Growth Centre, Knowledge for Trust Fund, MacArthur Foundation, Ewing Marion Kauffman Foundation, Social Science Research Council, Templeton Foundation, Earhart Foundation, Institute for Humane Studies, National Institute of Child Health and Human Development, National Institutes of Health, and National Science Foundation.

I was Coordinating PI for a € 13 300 000 European Research Council Synergy grant proposal "Difference-in-Indifference: Normative Commitments in Multiculturalist Societies" that advanced to the second stage in 2018 and Lead PI for a € 3 600 000 European Research Council Advanced grant proposal "E-Justice Innovations in the Wake of COVID-19" that advanced to the second stage in 2022.

I serve or have served on the Program Committees of NAACL Natural Legal Language Processing, International Conference on AI and Law, Econometric Society Meetings, European Economic Association, American Law and Economics Association, and European Law and Economics Association, and been invited to deliver keynotes at the European Law and Economics Association, French Law and Economics Association, Asia Law and Economics Association, International Conference on Computational Social Science (IC2S2), AI, law, and behavioral science conferences, and the 2018 Heremans Lectures in Law & Economics.

I am also lead PI (Principal Investigator) for DE JURE (Data and Evidence for Justice Reform), whose aim is to revolutionize how legitimacy and equality in justice systems are measured, understood, and enhanced. The goal is to move from studying historical data to working with administrative data, machine learning, and RCTs to achieve a more just system. The program has thus far worked with countries in three broad categories. In the first group, DE JURE works closely with court management, judiciaries, and training academies to design, deploy, and evaluate interventions—often developing the technologies to do so. In the second group, DE JURE works with auxiliary actors involved in access to justice to assess the effects and ability of trustworthy Artificial Intelligence (AI) to assess trust in the law. In the third group, DE JURE obtains data and conducts historical analyses on judicial efficiency or inconsistencies that may spur a cycle of change.

2 Background

I completed a MIT Economics PhD in 2004 under Esther Duflo—a trailblazer in using experimental evidence for economic policymaking—and a Harvard Law JD in 2009. After a tenure-track assistant professorship in law, economics, and public policy at Duke University (USA), I received tenure at ETH Zurich (Switzerland) in 2012 and co-founded the ETH Center for Law and Economics, where I held the Chair of Law and Economics with endowed funding of € 660 000 per year and start-up of € 580 000. In 2015, I came to Toulouse School of Economics (France) and its Institute for Advanced Study, which had five themes related to my work—law and economics, political

economy, behavioral economics, economic history, and economics and biology. In 2018, I was appointed Directeur de Recherche at the Centre National de la Recherche Scientifique (CNRS). In 2019, I was designated lead Principal Investigator for DE JURE (Data and Evidence for Justice Reform) at the World Bank.

Much of my work uses economics methodology to study legal concepts. Most of my publications are in economics journals (e.g., *American Economic Review*, *Quarterly Journal of Economics*, *Journal of Political Economy*, and *Econometrica*). Some articles are accepted at leading AI outlets including the flagship journal for AI and law, *Journal of Artificial Intelligence and the Law*. Some articles are in leading law journals and law and economics journals, like *Journal of Legal Studies*, *Journal of Law, Economics, and Organizations*, and *Journal of Law and Economics*.

In addition to my appointments at Toulouse, I am a regular visitor at NBER, project advisor at the NYU Courant Institute of Mathematical Sciences Center for Data Science, director of oTree Open Source Research Foundation (an open-science CERN for social science experiments), and founder of Data Science Justice Collaboratory (inspired by the Broad Institute, but for the legal domain). I have delivered over 700 presentations at universities, peer-reviewed conferences, and government venues, given the 2018 Heremans Lectures in Law & Economics at KU Leuven, and instructed at spring/summer schools in law and economics, organizational, and behavioral economics.

3 Teaching

My core teaching philosophy is one of empathy and crafting an interactive lesson plan that sparks curiosity. My economics education engendered a scientific approach focusing on “why” and “how do we know”. My legal education imparts a soft-socratic method so that students arrive at answers for themselves. Every day is an exercise in empathy – listening – and sparking curiosity with a soft-socratic method. This approach likewise infuses my efforts to give students the tools to answer questions on their own and to ask new questions. Providing such tools is what inspired a project, oTree, which can be used by college economics students as part of a class on game theory or social psychology or for high school students who want to learn Python. My aim was to increase recognition and respect across cultural boundaries by giving students the tools to understand better each other’s revealed preferences. Through these teaching, mentoring, and outreach efforts, I have been recognized with nominations for the Joseph R. Levenson Memorial Teaching Prize (2009) and John H. Marquand Award for Exceptional Advising and Counseling (2008), both at Harvard College.

My past teaching experience also included developing new courses: (1) Theorizing Cultural Differences-The Economics of Fundamentalism (at the University of Chicago), (2) Hermemetrics: The Economics of Interpretation and (3) Hermemetrics Lab (both at Harvard College). I have also taught (4) Contracts Law (1L) at Duke Law School. Beyond this, I have given the spring summer school lectures on (5) Positive Foundations of Normative Commitments at the Institutional and Organizational Economics Academy, a graduate course on (6) Experimental Economics: Sources of Normativity at Toulouse School of Economics, Lectures in Law and Economics on (7) Legitimacy, Law, and Recognition-Respect at KU Leuven, and (8) law and economics (theoretical and empirical) for law students and graduate students at Duke University and Toulouse School of Economics. In addition, I have served as (9) project advisor for machine learning and computational statistics classes at NYU Courant Institute of Mathematics Center for Data Science and taught on the topic at Toulouse School of Economics. I have also taught (10) executive education for Toulouse School of Economics. Through my work at DE JURE (Data and Evidence for Judicial Reform), I have had to teach concepts to general audiences for a wide range of international professionals.

4 Outline

The development of institutions is critical to preventing tragedies that stem from lack of choice, be the tragedies ones of violence, poverty, infectious diseases, or lack of health and education. But incentives to help members of shared identities frequently debilitate these institutions. On a thick vein, my work studies the macro and demographic forces and laws that aid or hinder discrimination in legal institutions, markets, and public policy. My work in law and development traces the incentives that lead to human rights violations, such as religious, ideological, and gender-based violence. My early work examines how acts of violence derived their legitimacy from proto-legal sources and ultimately undermined that legitimacy. In **theorizing cultural differences: the economics of fundamentalism**, I document economic forces underlying the religious provision of social insurance, social sanctions, and social conservatism and, turning to one dimension of social conservatism, the economic incentives that give rise to gender violence, sexual harassment, and the regulation of the private domain. In doing so, my work offers a theory for why and where church-state separation arose and develops high-dimensional methodologies for **measuring the consequences of normative commitments** using the random assignment of judges.

My work on markets and morality draws on experience with randomized control trials in developing countries to conduct a series of experiments in institutional and mechanism design that blur lab and field with the use of disaggregated labor markets. Observational data offers one lens on human behavior with which to evaluate legal rules. Experimental data offers another, but manipulating legal rules in field settings is practically infeasible. Motivated by the question, **do free markets corrode moral values**, I test behavioral assumptions underlying the perceived dichotomy between law and economics approaches to law—morality vs. markets, deontological vs. consequentialist, expressive vs. incentives, subjective vs. objective, and duty vs. damages. On a theoretical level, **distinguishing sacred values from social preferences: theory, evidence, and relevance of deontological motivations** shows how the presence of these commitments can problematize empirical methods widely used in economics (i.e., the random lottery incentive and strategy method) and political science, sociology, and psychology (i.e., Likert scales). The presence of non-consequentialist motives challenge theoretical models that assume a convex—rather than concave—cost of deviating from moral and ideological bliss points. On a practical level, modularizing the code used for these studies led to the development of oTree—an open-source programming language for lab, online, and field experiments. oTree enhances the ability to study normative economics, empirical moral psychology, experimental philosophy, and interactive epistemology in contextualized settings. It also enables researchers to explore the transmission and persuasion process of normative commitments whose incommensurability can lead to conflict and violence. oTree was used in an EU Horizon 2020 project to study behavior in large-scale networks, voting, macroeconomics, and mixed agent-based experiments. I use oTree in several field experiments with government partners not feasible with alternative measurement tools: a study evaluating compulsory economics in France for a half-million students, evaluating social-emotional learning interventions and training of judges and prosecutors in Peru, nationwide experiment using quadratic voting for survey research in Estonia, empathy training experiment in the Civil Servant Academy in Pakistan, grit experiment for children amid lockdown in the United Arab Emirates, and a study on how youths learn justice via an embedded-ethics digital literacy curriculum in India and Tanzania.

My work on behavioral judging interrogates the **impossibility of objective judgments: priming, gambler’s fallacy, mood, voice, and peer effects in courts**. My work has digitally curated 12 terabytes of archival and administrative data on judges and courts where normative ideas incubate, textually and orally. I use these data to study the social and psychological, economic and political influences in **judicial analytics and the behavioral foundations of polarization**. Two singular initiatives are: 1) the digital universe for a century-and-a-half of roughly a half-million cases in the U.S. Courts of Appeals, engineered 2 billion N-grams up to eight words in length, and constructed the networks of citations across cases and seating among judges, who are randomly assigned to panels of three. These cases are linked to the universe of the U.S. Supreme and District Courts, Administrative Office of the U.S. Courts, and hand-collected features in a 5% random sample and 2) digitized speech formants in

2500 hours of U.S. Supreme Court oral arguments for over a half-century—longitudinal data on speech intonation (linguistic turns) are rare—and manually clipped oral advocates’ identical introductory sentences—speech variation holding words fixed is rare in high-stakes settings. I also curated the universe of 1 million decisions and 15 million hearing sessions in U.S. immigration courts, the universe of 1 million criminal sentencing decisions in U.S. District Courts, and the universe of judges’ biographies that are the potential source of their normative commitments. An additional 1000 legal databases at all levels of federal and state government are tagged and linked. Similar datasets are being constructed for India (8.7 million case records and 67 million hearings from 24 High Courts and 3000 subordinate courts), Kenya, Chile, Peru, and Indonesia. The initiative is inspired by the Broad Institute, but for the legal domain.

My work on law and legitimacy formalizes recognition-respect theory and what it means for legal institutions, lawmakers, judges—to be indifferent, such that it violates notions of justice. **Deter or Spur? British Executions During World War I** digitized and linked eleven World War I British archival datasets comprising the universe of deserters reported in military diaries, police gazettes, and handwritten military trials, commuted and executed capital sentences, geocoded casualties, maps, officer lists, and the military units’ order of battle. I documented differential effects of executing British and Irish soldiers on their subsequent desertions. Second, I introduce the idea of **difference in indifference: legitimacy, law, and recognition-respect**. I distinguish sympathy from empathy—the latter linked to recognition, knowledge and respect of reference points. Information acquisition can be endogenous to preferences, which blurs the boundary between Becker’s statistical and taste-based discrimination. Psychologists find many effects of moderate sizes in the lab, so settings where people are closer to indifference among options are more likely to lead to detectable effects outside of it. As preferences over the legally relevant factors wanes, the influence of extraneous factors grows. Measuring behavioral biases is one way to document revealed preference indifference. Third, I illustrate how machine learning can help detect indifference using data on asylum judges and criminal prosecutors (who handle fifteen times the number of cases that judges do, but whose role is largely with little accountability). Early predictability suggests the possibility of snap or pre-determined judgments, which renders the possibility that algorithms may be less biased than human decision-makers. I demonstrate using machine learning to detect characteristics noisy to human decision-makers with a unique dataset recording a half million defendants for a decade, vertically linked from the police arrest to the potential final sentence (vertical linkages from the time of arrest, including those sent home without a trial, otherwise do not exist). Fourth, I introduce an incremental approach to AI aiding normative decision-making, which leverages self-image motives to facilitate its adoption, to be tested via randomized control trials with prosecutors in Brazil.

My work on the demography of ideas traces the **genealogy of ideology** in the corpus of judicial decisions. I investigate the impact of corporate sponsored economics training on moral reasoning and criminal justice (sentencing harshness and disparate impacts become more justifiable under theories of deterrence and statistical discrimination). In parallel work, I use micro-level data on physician decision-making to investigate the effects of pharmaceutical company payments to physicians on their prescriptions, patient outcomes, and patient adherence. I link administrative, unencrypted (through special permission) Medicare data to industry-physician relationships cleaned from litigation settlements prior to the Affordable Care Act and from the universe of payments reported after the Affordable Care Act and linked to NDC drug-code—a combined total of roughly 30 million payments. I also investigate the link between these payments to malpractice, physician inattention, and the p-curve of underlying scientific papers (linked by NDC drug code) to trace the lives cost by the scientific reproducibility crisis.

My work has broadly been motivated by an idea that can be termed, **hermemetrics: the measurement of meaning**, which is hermeneutics and econometrics or economics of interpretation, as a source of normative commitments. Hermemetrics examples include: (a) modeling the economic incentives behind the shift from pro-welfare religious interpretations a century ago to the anti-welfare posture of today’s religious right and why church-state separation arose when and where it did, (b) measuring the effects of sexual harassment law—an interpretation of

anti-discrimination law—on gender inequality by exploiting the random assignment of appellate judges interpreting the facts and the law differently but in a statistically predictable manner, (c) using a particular instance of interpretive injustice where British capital cases during World War I were randomly executed or commuted to estimate the deterrent and delegitimizing effects of the death penalty, (d) studying how individuals misinterpret the law by manipulating tax schedules and contracts in labor market settings, (e) understanding why and when lay interpretation of legal decisions have expressive or backlash effects, and (f) revelation of deontic motivations through revealed preference and in interpretations of legal and social texts. If economics is the science of how individuals optimize to budget constraints, hermometrics studies how societies optimize in response to textual constraints.

Most of the papers and slides can be directly accessed at: <http://users.nber.org/~dlchen/papers/>

5 Law and Development

In what follows, I present abstracts of my past work organized by theme—I like to think about the themes in terms of book projects. Submissions have been requested from *Harvard (2)*, *Oxford (2)*, *Yale*, *MIT*, *Virginia*, and *Palgrave*. Due to space constraints, the abstracts can be clicked to link to the paper or book-length manuscripts online.

5.1 Theorizing Cultural Differences: The Economics of Fundamentalism

One of the most influential views of our time attributes a large part of the failure of development in the post-war period to group conflicts. Recent research in development economics has identified a large collection of policy innovations that would help the poor. But these policies often do not get adopted because of conflicts between groups. My Ph.D. research at MIT examined the economics of fundamentalism.

1. Club Goods and Group Identity: Evidence from Islamic Resurgence During the Indonesian Financial Crisis (*Journal of Political Economy*, 118(2), 300-354, 2010) tests a model in which group identity in the form of religious intensity functions as ex post insurance. I exploit relative price shocks induced by the Indonesian financial crisis to demonstrate a causal relationship between economic distress and religious intensity (Koran study and Islamic school attendance) that is weaker for other forms of group identity. Consistent with ex post insurance, credit availability reduces the effect of economic distress on religious intensity, religious intensity alleviates credit constraints, and religious institutions smooth consumption shocks across households and within households, particularly for those who were less religious before the crisis.
2. Islamic Resurgence and Social Violence During the Indonesian Financial Crisis (*Institutions and Norms in Economic Development*, MIT Press, ed. M. Gradstein and K. Konrad, 179-200, 2007) also focuses on the intensity with which people identify with their groups. Violence is a negative externality with enormous social costs, so to the extent group identity and social violence (physical acts of destruction, killing, looting, attacks, burning, clashes, taking hostages, etc., by one group against another) are related, policies taking into account intensity of group identity need to be considered. This paper documents a correlation between group identity and group conflict in the specific context of Islamic resurgence during the Indonesian financial crisis.
3. Economic Distress Stimulates Religious Fundamentalism presents a causal analysis. It exploits relative price shocks induced by the 1997 Indonesian financial crisis and variation in religious institutions across Indonesia before the crisis to identify the effect of economic distress on the relationship between religious institutions and social violence. In the cross-section, high religious intensity areas before the crisis have more social violence after the crisis. In the panel, social violence increases fastest where participation in Koran study also increases the fastest. In the two-stage least squares, instrumenting for economic distress using relative price

shocks shows a causal relationship between economic distress and the relationship between religious intensity and social violence. These results are unlikely to be driven by omitted environmental variables. Credit availability mitigates this effect. Economic distress alone did not stimulate social violence but stimulates it in the presence of religious institutions. I explain these findings in a model where high marginal utilities during economic distress increase incentives for group conflict where group conflict increases the budget of insurance groups. With volatility, religions with stronger sanctions or more out-group bias are more stable and successful. As volatility declines, benign groups and religions become relatively successful.

4. Gender Violence and the Price of Virginty: Theory and Evidence of Incomplete Marriage Contracts (*Journal of Religion and Demography*, 2020) builds and tests a model of marriage as an incomplete contract that arises from asymmetric virginity premiums and examines whether this can lead to social inefficiencies. Contrary to the efficient households hypothesis, women cannot prevent being appropriated by men once they enter marriage if they command lower marriage market opportunities upon divorce. Because men cannot or do not commit to compensating women for their lower ex post marriage market opportunities, marriage is an incomplete contract. Men may seek to lower women's ex ante "market wages" in order to induce entry into joint production. Inefficient or abusive marriages are less likely to separate. Equalizing virginity premiums may reduce domestic and non-domestic violence. Female circumcision and prices women pay doctors to appear virgin before marriage in many countries suggest asymmetric virginity premiums continue to exist. Evidence from China and the U.S. suggest asymmetric virginity premiums persist over economic development. Asymmetric virginity premiums are strongly positively correlated with female but not male virginity premiums. I use variation in religious upbringing to help estimate the effect of virginity premiums on gender violence in the U.S. The OLS relationship between virginity premiums and female reports of forced sex may be biased downwards if shame is associated with abuse and this shame is greater for women with higher virginity premiums. But the OLS relationship for males might not be biased downwards. Asymmetric virginity premiums are positively correlated with men forcing sex on women and paying women for sex. The model complements a growing empirical literature on inefficient households and human rights abuses, visible manifestations of female appropriability across time and space.
5. Why are religious groups with greater within-group charitable giving more socially conservative and opposed to the welfare state? The Political Economy of Beliefs: Why Fiscal and Social Conservatives/Liberals Come Hand-in-Hand (*Journal of Political Economy: Micro*; under review, J. Lind) proposes and tests a theory where religious provision of social insurance explains why fiscal and social conservatism align. Religious groups that provide social insurance act as competitors to state-run insurance, and therefore resist expansion of the state into the provision of social insurance. The alignment between fiscal and social conservatism disappears when there is a state church and reverses for members of a state church. Moreover, plausibly exogenous increases in church-state separation precede increases in the alignment between fiscal and social conservatism, which mitigates concerns of the reversal being due to omitted environmental variables. The theory provides a novel explanation for religious history: as elites gain access to alternative social insurance, they judicate increasing church-state separation to create a constituency for lower taxes. This holds if religious voters exceed non-religious voters, otherwise, elites prefer less church-state separation in order to curb the secular left, generating multiple steady states where some countries sustain high church-state separation, high religiosity, and low welfare state, and vice versa. Religion, Welfare Politics, and Church-State Separation (*Journal of Ecumenical Studies*, 42(1), 42-52, 2007; J. Lind) uses this framework to explain the changing nature of religious movements, from Social Gospel to the Religious Right, and why church-state separation arose in the U.S. but not in many European countries.

5.2 Measuring the Consequences of Normative Commitments

To address human rights violations, societies often turn towards laws, upheld and interpreted by judges. When there is no strong legal precedent, the current method of judicial decision-making relies on informal policy analysis. Interpretive discretion opens the possibility of bias. Recent econometric advances make possible methodologies for empirically evaluating the effects of law using the random assignment of judges interpreting the facts and the law in different ways. Measuring the Moral and Economic Consequences of Judicial Discretion is a book-length treatment, which proposes rapid impact analyses to aid judges deliberating hard cases.

My J.D. research at Harvard Law first developed this method to study the impact of sexual harassment law. Feminist legal theory introduced the concept of sexual harassment into American jurisprudence in the 1970s; over the next few decades, sex discrimination laws were reinterpreted to include sexual harassment. Subsequent papers extend to other legal domains, address theoretical questions, buttress with experimental evidence, and formalize econometrically a framework with broad applicability as well as assess assumptions underlying the methodology—state and judicial compliance, awareness, area effects, exclusion restriction, and low-dimensionality.

1. Is labor market gender inequality due to physiological differences, labor market choices, or discrimination? *Insiders, Outsiders, and Involuntary Unemployment: Sexual Harassment Exacerbates Gender Inequality (Management Science, reject & resubmit; J. Sethi)* uses novel data on all workplace sexual harassment appellate precedent from 1982-2002 and randomly assigned judges. We find that pro-plaintiff sexual harassment precedent reduced gender inequality and spurred the adoption of sexual harassment human resources policies. The effects were comparable to the Equal Employment Opportunity Act’s impact on black employment share, greatest in the heavily-litigated construction industry and where male sexism was highest, and explains one-third of the adoption of sexual harassment human resources policies. Our results are consistent with an insider-outsider theory of involuntary unemployment.
2. *Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain (Econometrica, 80(6), 2369-2429, 2012; A. Belloni, V. Chernozhukov, C. Hansen)* develops results for the use of Lasso and post-Lasso methods to form first-stage predictions and estimate optimal instruments in linear instrumental variables (IV) models with many instruments, p . Our results apply even when p is much larger than the sample size, n . We show that the IV estimator based on using Lasso or post-Lasso in the first stage is root- n consistent and asymptotically normal when the first stage is approximately sparse, that is, when the conditional expectation of the endogenous variables given the instruments can be well-approximated by a relatively small set of variables whose identities may be unknown. We also show that the estimator is semi-parametrically efficient when the structural error is homoscedastic. Notably, our results allow for imperfect model selection, and do not rely upon the unrealistic “beta-min” conditions that are widely used to establish validity of inference following model selection. In simulation experiments, the Lasso-based IV estimator with a data-driven penalty performs well compared to recently advocated many-instrument robust procedures. In an empirical example dealing with the effect of judicial eminent domain decisions on economic outcomes, the Lasso-based IV estimator outperforms an intuitive benchmark. Optimal instruments are conditional expectations. In developing the IV results, we establish a series of new results for Lasso and post-Lasso estimators of nonparametric conditional expectation functions which are of independent theoretical and practical interest. We construct a modification of Lasso designed to deal with non-Gaussian, heteroscedastic disturbances that uses a data-weighted l1-penalty function. By innovatively using moderate deviation theory for self-normalized sums, we provide convergence rates for the resulting Lasso and post-Lasso estimators that are as sharp as the corresponding rates in the homoscedastic Gaussian case under the condition that $\log p = o(n1/3)$. We also provide a data-driven method for choosing the penalty level that must be specified in obtaining Lasso and post-Lasso estimates and establish its asymptotic validity under non-Gaussian, heteroscedastic disturbances.

3. Mostly Harmless Machine Learning: Learning Optimal Instruments in Linear IV Models (*NeurIPS 2020 Workshop on Machine Learning for Economic Policy*; J. Chen, G. Lewis) provide some simple theoretical results that justify incorporating machine learning in a standard linear instrumental variable setting, prevalent in empirical research in economics. Machine learning techniques, combined with sample-splitting, extract nonlinear variation in the instrument that may dramatically improve estimation precision and robustness by boosting instrument strength. The analysis is straightforward in the absence of covariates. The presence of linearly included exogenous covariates complicates identification, as the researcher would like to prevent nonlinearities in the covariates from providing the identifying variation. Our procedure can be effectively adapted to account for this complication, based on an argument by Chamberlain (1992). Our method preserves standard intuitions and interpretations of linear instrumental variable methods and provides a simple, user-friendly upgrade to the applied economics toolbox. We illustrate our method with an example in law and criminal justice, examining the causal effect of appellate court reversals on district court sentencing decisions.
4. Is it justified for states to appropriate private property rights? If so, should governments expropriate or regulate? Government Expropriation Increases Economic Growth and Racial Inequality: Evidence from Eminent Domain (*JPE Macro*, under review; S. Yeh) tests three conventional views: insecure property rights cause underinvestment, moral hazard cause over-investment, or public use cause economic growth. We embed these mechanisms in a model and measure them using the random assignment of U.S. federal court judges setting geographically-local precedent. For a half-century, racial minority Democrats were more likely to strike down government appropriations while Republican former federal prosecutors were more likely to uphold them. We find that pro-government physical takings precedent stimulated subsequent takings, expropriation of larger parcels, highway construction, and growth in construction, transportation, and government sectors as well as agriculture, retail, and financial sectors, overall economic growth, and property values. However, racial minorities were increasingly displaced, unemployed, and living in public housing, and the service sector declined. Pro-government regulatory takings precedent also spurred economic growth and property values, but did not increase displacement or racial inequality.
5. Whether policies shift preferences is relevant to policy design. How Do Rights Revolutions Occur? Free Speech and the First Amendment (*Economic Journal*, revise & resubmit, 2019; S. Yeh) tests a model of law and norms using an area of law where economic incentives are arguably not the prime drivers of social change. From 1958–2008, Democratic judges were more likely than Republicans to favor progressive free speech standards. Using the random assignment of U.S. federal court judges setting geographically-local precedent, we estimate that progressive free speech standards liberalized sexual attitudes and behaviors and increased both crime rates and the spread of sexually transmitted diseases. We then randomly allocated data entry workers to enter news-articles of court decisions. Progressive decisions liberalized sexual attitudes and shifted norm perceptions even for data entry subjects, but not self-reported behavior. These results present evidence of law’s expressive power.
6. To build on the previous study and the field data using random assignment of judges, The Construction of Morals (*Journal of Economic Behavior and Organization*, 104, 84-105, 2014; S. Yeh) reports contextualized field experiments. We randomize data entry workers to transcribe newspaper summaries of liberal or conservative court decisions about obscenity. As before, we find that liberal obscenity decisions liberalize individual and perceived community standards and increase utility. Yet we also find that religious workers become more conservative in their values, identify as more Republican, view community standards as becoming more liberal, and report lower utility. Workers update beliefs about the prevalence of sexual activities differently in response to liberal or conservative decisions. These results provide causal evidence for the law having indirect social effects that may amplify or attenuate deterrence effects and suggest that legitimacy of law can affect

utility and self-identification.

7. The previous framework relies on a model of law having backlash *or* expressive effects, but the next paper develops a theoretical model where law has backlash, *then* expressive effects. Turning to courts to vindicate rights often led to resistance and subsequent acceptance. Can Policies Affect Preferences? Theory and Evidence from Random Variation in Abortion Jurisprudence (*American Economic Journal: Economic Policy*, revise & resubmit; V. Levonyan, S. Yeh) develop a model where laws reducing costs of acts that diminish self-image generate temporary backlash. We present a two-layered judge randomization framework for estimating effects of presence and direction of court rulings. Randomly-assigned judges' politics, religion, and ethnicity predict abortion rulings, which affect state regulations previously found to restrict abortions and impact society. Consistent with theory, stated and revealed preference (3 million donations) shifted against legalized abortion, especially among Republicans and on discretionary abortions, only in short-run, echoing historical analyses and an original data entry experiment.
8. Just Water? Environmental Jurisprudence, Water Quality, and Infant Mortality in India (S. Bhupatiraju, S. Joshi, P. Neis, S. Singh) Toxic ambient water can be deadly, particularly in developing countries where pollution levels are high. The failure of executive and legislative policies to address water pollution across the world have often prompted judiciaries to respond to the challenge of protecting ambient water. We explore the impacts of judicial policies on surface water quality in the case of India. We construct a novel dataset that combines court cases, histories of the judges who ruled, surface-water pollution, child mortality and socio-economic development at the level of districts. We use an instrumental variables framework to exploit the process of random judge assignment within the Indian justice system to estimate the effect of "green" rulings on district-level pollution in India. We find that the rulings precipitated immediate reductions in surface water pollution, particularly in the Ganga river basin. These effects however, are confined to the year of the ruling and do not show persistence. We also find limited impacts on neonatal and infant mortality. This is suggestive evidence that judicial policies can succeed in lowering short-term pollution but the broader problems of environmental governance limit the power of judiciaries in environmental protection.
9. Do Judicial Sentiments Affect Social Attitudes? (*Economica*, 89, 362–376, 2022; E. Ash, S. Galletta) asks the question, do judge's attitudes, reflected in court opinions, impact population attitudes? We present an application of machine learning and causal inference using the universe of U.S. Circuit Court opinions from 1891. We assess the sentiment of each paragraph on thermometer (sentiment) questions in the American National Election Survey. We calculate cosine similarity of each paragraph to each thermometer target (e.g., Republicans, Democrats, woman, feminists, etc.) using word2vec, and use a sentiment analyzer to compute the average sentiment (positive or negative) of each opinion towards each target. We then use LASSO to select biographical characteristics predictive of attitudes and use the selected characteristics as instruments in a two-stage least square regression. We show that sentiment reflected in court opinions impact population attitudes in an opposite direction. These results complement other quantitative and qualitative work suggesting that court actions often led to resistance.
10. Law and Norms: A Machine Learning Approach to Predicting Attitudes Towards Abortion (K. Kwan, M. Maass, L. Ortiz) uses U.S. Courts of Appeals cases, sociological and attitudinal indicators, criminal statistics, and a variety of survey data to predict societal attitudes towards abortion. Understanding the predictors of societal attitudes has been widely investigated using individuals' responses to surveys and polls. We create two classification models: pro-abortion attitudes for health reasons related to the mother, fetus, or rape and pro-abortion attitudes for any other personal non-health related reason. To address high dimensionality, we employ factor analysis to group indicators. Logistic regression and random forests performed best among

three types of classifiers evaluated on AUC and accuracy. For pro abortion attitudes related to health related abortions, the most important factor contains sociological and attitudinal indicators about the frequency of contact with family and friends. The most important U.S. Courts of Appeals indicators include the religion, political affiliation, and ABA ratings of the judges in the Circuit pool. The most important crime indicators are the rates of violent and property crimes. For pro abortion attitudes towards non-health related abortions, the most important factor contains sociological and attitudinal indicators about job satisfaction, religious preference, religion raised in, and beliefs about the Bible. The most important U.S. Courts of Appeals indicators include the religion and political affiliation of the judges in the Circuit panel. These legal indicators are grouped in the same factor as the most important crime indicators, which are rates of crimes against society, crimes against property, and violent crimes.

11. Distinguishing Between Custom and Law: Empirical Examples of Endogeneity in Property and First Amendment Precedents (*William & Mary Bill of Rights Journal*, 21(1081), 2013, with S. Yeh) discusses the relationship between custom and law to highlight the phenomenon of endogeneity when empirically evaluating the effects of laws. An important literature evaluates the roles of laws in motivating behaviors, including investigations of whether or how laws influence customs and social norms. Traditional economic analysis, for example, posits that official or codified laws are expected to influence behaviors by formally incentivizing a particular action or communicating values. Enhancing this strand of thought, there is a growing contribution from historical and empirical analyses that link laws to broader societal changes over time. At the same time, a valuable discourse examines how customs may determine both de facto laws and formally enacted laws, including the court precedents that are rendered. Whether they are directly codified into a legal test or informally referenced, customs can influence the formal laws that are adopted in a community and beyond. Indeed, some scholars have argued that evolving customs and norms have influenced the Court in its decisions. The subsequent effects of these formal laws and court decisions are of tremendous interest to policymakers and judges, and with policy concerns in mind, we argue that one must not ignore the endogenous feedback between aggregate behaviors, customs, and laws. That is, while customs may shape or influence laws, laws can also shape customs through their effects on behaviors or norms in the aggregate. The endogeneity that custom produces suggests that simply by observing a correlation between law and behavior is not enough to assert that a law in itself is effective or to assert that social trends and evolving customs are driving legal change. We propose a greater role for experimental or quasi-experimental evidence among scholars interested in the empirical study of law and sketch out an original empirical strategy that could overcome the endogeneity between custom and law,
12. Do judges directly affect economic outcomes at the time they are revealed? The Shareholder Wealth Effects of Delaware Litigation (*American Law and Economics Review*, 19(2), 287-326, 2017; A. Badawi) collects data on the record of every action in hundreds of derivative cases and merger class actions involving public companies filed in the Delaware Court of Chancery from 2004 to 2011. We use these data to estimate how markets respond to litigation in the most important court for corporate disputes in the United States. The detail in the dataset allows us to explore how case characteristics such as the timing of case filing, the presence of certain procedural motions, litigation intensity, and judge identity relate to firm value. Unlike previous studies, we document that that negative abnormal returns are associated with the filing of derivative cases, and that association is particularly strong for cases that are first filed in Delaware and are not related to a previously disclosed government investigation. We also develop some evidence that market participants can anticipate litigation intensity and respond by valuing the firm equity less. Finally, we find little evidence of abnormal returns associated with judicial assignment at the time of filing of derivative cases, but we do observe an association between judicial assignment and case filing for merger cases.

13. Do people pay attention to court outcomes? Does Appellate Precedent Matter? Stock Price Responses to Appellate Court Decisions of FCC Actions (in Empirical Legal Analysis: Assessing the Performance of Legal Institutions, 2013, with S. Yeh and A. Araiza) tests the effects of federal appellate court decisions of Federal Communications Commission (FCC) actions on stock prices using differences-in-differences (DID) and an instrumental variables approach. This study exploits the random assignment of appellate judges to three-judge panels and the fact that a judge's (1) party affiliation and (2) race predict outcomes in appellate court decisions of FCC actions to instrument for anti-industry decisions, which favor the public interest. This instrumental variable approach demonstrates a causal relationship between appellate court decisions of FCC actions and changes in stock prices of media firms relative to the stock prices of non-media firms. The differences-in-differences (DID) analysis shows that federal appellate court decisions against media businesses decreased media stock prices. The instrumental variables analysis shows that these appellate court decisions decreased media stock prices relative to non-media stock prices, one and two years after the court decisions. Recent studies indicate that stock prices serve as a proxy for competition and that decreased media competition may correspond to an increase in variety of programming. These findings suggest that when deciding against media businesses, the courts effectively reinforced the purpose of the FCC to serve the public interest by promoting a diversity of viewpoints.
14. Do judges follow legal precedent? Public enforcement of law relies on the use of public agents, such as judges, to follow the law. Are judges motivated only by strategic interests and ideology, as many models posit, rather than a duty to follow the law? Judicial Compliance in District Courts (*International Review of Law and Economics*, forthcoming) uses the random assignment of U.S. Federal judges setting geographically-local precedent to document the causal impact of court decisions in a hierarchical legal system. We examine lower court cases filed before and resolved after higher court decisions and find that lower courts are 29-37% points more likely to rule in the manner of the higher court. The results obtain when the higher court case was decided in the same doctrinal area as the pending case and when the higher court case was decided on the merits. Reversals by the higher court have no significant effects. These results provide clean evidence that judges are motivated to follow the law and are not solely motivated by policy preferences.
15. Religious Freedoms, Church-State Separation, and Religiosity: Evidence from Randomly Assigned Judges provides evidence on how laws related to the establishment or exercise of religion affect the level of religious activity. We exploit the random assignment of judges to regional appellate courts, along with the fact that judge politics, religion, and other characteristics predict decision-making in church-state jurisprudence, to obtain exogenous variation in religion caselaw. Our first-stage effect is driven by the fact that Democratic, minority-religion judges tend to favor greater separation of church and state than their Republican, majority-religion counterparts. In the second-stage estimates for the effect of law on outcomes, we report two findings. First, judicial decisions reducing government subsidies to mainstream religions reduce measures of supply-side religious services. Second, judicial decisions that strengthen the free exercise of fringe religions increase self-reported religiosity.
16. Do stronger constitutional assembly protections increase the number of political protests? Do those protests result in more democratic political processes? Protest and Political Accountability: The Electoral Effects of Protest Rights and Rates (E. Reinhart) looks at the durable impacts of long-term differences across jurisdictions in protests. We find that protest rights are more likely to be upheld by judges who used to work in local courts, federal district courts, or local city councils. A pro-protest circuit decision increases the number of protests by 7%. Doubling the number of protests is associated with an 11% increase in the incumbent vote share within-state.
17. Does Dicta Matter? Evidence from Environmental Law poses the question whether judicial decisions affect

outcomes only through their verdicts alone, or whether the manner of an opinion’s writing (dicta) also affects outcomes. We characterize U.S. federal judges as vectors of their opinions and analyze the causal impact of random assignment of judges with different opinion styles in environmental law. We document a very strong first stage in opinion style, for example, persistent citation habits. Using cross-validated prediction of judicial rulings, we show pro-EPA rulings reduce air pollution.

18. Deep IV in Law: Automated Impact Analysis of Court Precedent and Application to Criminal Sentencing (in *Frontiers in Law*, Cambridge University Press, ed. A. Bindler and R. Ferrer; Z. Huang, R. Wang, X. Zhang) explores the use of high-dimensional instrumental variables to estimate the causal relationship between criminal appeals in U.S. Circuit Courts and lengths of criminal sentences imposed by U.S. District Courts. Using judge characteristics as instruments, we implement two stage models on court sentencing data for the years 1991 through 2013. We find that Democratic, Jewish judges tend to be in favor of criminal defendants, while Republican, Catholic judges tend to rule against criminal defendant. However, there is weak evidence that a high-dimensional representation of the court ruling yields causal impacts. Methodologically, we demonstrate the applicability of deep instrumental variables to legal data.

5.3 The Role of Justice in Development

1. The Role of Justice in Development: The Data Revolution (*Journal of Economic Perspectives*, under review; M. Ramos-Maqueda) Efficient, fair, and accessible justice systems promote peace and security, support economic investment and growth, and provide fundamental protections to citizens essential for sustainable poverty reduction and increased shared prosperity. This article summarizes the evidence on justice for economic development, conflict, and corruption. It then considers the promise of administrative data, machine learning, and randomized control trials to enhance efficiency, access, and quality of justice and evaluates the potential for e-justice built on the data revolution in the age of COVID-19. A version of this argument focused on the Middle East was published as Transforming Justice in the Middle East and North Africa through Data (Middle East and Near Africa World Bank Flagship Report; M. Ramos-Maqueda). We highlight the importance of data for the judiciary of India here: A Decade Later, POSCO Isn’t Perfect, But It’s Had an Impact (*Times of India*, Nov 21, 2022; A. Ranjan, S. Bhupatiraju, S. Joshi).

6 Markets and Morality

6.1 Do Free Markets Corrode Moral Values?

My early economics research studied how market forces *interact* with normative commitments. My later legal research studied the impact of market forces *on* normative commitments. At least since Adam Smith and David Hume, scholars have offered hypotheses about the effect of a citizen’s economic experience on his or her moral life as an individual. It has been asserted that competition may bring a winner-take-all mentality and a lack of concern for others or that exposure to market values will lead us to abandon non-utilitarian forms of moral thought, treating every moral issue instead in terms of costs and benefits. (Whether this is viewed as a bad thing, of course, depends on one’s attitude towards utilitarianism.) On the positive side, the proponents of the so-called *doux commerce* thesis (a theory popularized by 18th century political philosophers) have proposed that a competitive market, with its disruptive effect on geographical and tribal isolation, will actually have morally improving effects, increasing our care for and understanding of others. Of particular interest today are questions like: Could the structure of employment affect moral attitudes? What is the effect of pharmaceutical company payments to physicians, and the potential role for regulation, transparency, and accountability in medicine? What is the impact of third-party litigation funding? What is the impact of market forces on law?

1. Markets and Morality: Do Free Markets Corrode Moral Values? (*Journal of Law, Economics, and Organization*, conditionally accepted; E. Reinhart) studies the impact of employment structure on three normative issues: utilitarian versus deontological values, other-regarding preferences, and charitable donations. Through a labor market intermediary, I randomly assigned workers to competitive or piece-rate work conditions. The groups were given a moral question posing a conflict between utilitarian and deontological values, and offered a choice to make a charitable donation. The moral question was accompanied by an illustration that made salient out-group considerations. Several results emerge: Competitively structured work experiences increased deontological value choices, deontological commitments towards out-group members, and donations by productive workers relative to non-productive workers; and the effects on deontological value choices differ by a country's level of economic development. I reconcile these results with a theoretical model based on experimental findings in affective moral psychology. When competition is perceived as unfair or unfamiliar, negative affect triggers deontological value choices, but when it is perceived as familiar or even fun, positive affect increases utilitarian attitudes. If utilitarian attitudes lead to market-oriented policies, multiple steady states arise where some countries sustain high utilitarianism, market-orientation, and economic growth, and vice versa. I use this perspective to help explain the intellectual history of the *doux commerce* thesis.
2. Do Markets Overcome Repugnance? Muslim Trade Response to Anti-Muhammad Cartoons (*European Economic Review*, revise & resubmit) tests a theory of trade, trust, and conflict. After Danish newspapers published Anti-Muhammad cartoons—on the intensive margin—exports from Danish countries to Muslim countries decreased by 23% for 20 months but rebounded by 24% thereafter. More secular Muslim countries were less affected by the controversy. Final goods (food, transport, and manufactured goods) were more affected, while intermediate goods (crude materials and commodities exports) were unaffected. The rebound virtually eliminated Muslim backlash to the obscenity deemed blasphemous from an Islamic perspective. Interestingly, Islamic exports to Denmark were unaffected—Danish money was acceptable—throughout this time period. On the extensive margin—the number of product categories trading—did not rebound. These results present a mixed picture on the costs of free speech. They are consistent with trade models with fixed costs preventing importers from easily shifting between countries, and consistent with markets sometimes eliminating the effect of moral attitudes against opposing ideologies.
3. Mandatory Disclosure: Theory and Evidence from Industry-Physician Relationships (*Journal of Legal Studies*, 48(2), 409-440, 2019; V. Levonyan, E. Reinhart, G. Taksler) presents and test a model of mandatory disclosure. The effects of disclosure laws on what is being disclosed are typically unknown since data on disclosed activity rarely exist in the absence of disclosure laws. We exploit data from legal settlements disclosing \$316 million in pharmaceutical company payments to 316,622 physicians across the U.S. from 2009-2011. States were classified as having strong, weak, or no disclosure based on whether the data was reported only to state authorities (weak) or were publicly available (strong). Strong disclosure law was associated with reduced payments among doctors accepting less than \$100 and increased payments among doctors accepting greater than \$100. Weak disclosure states, despite imposing administrative compliance costs to industry, were indistinguishable from no disclosure states. This result suggests that the mechanism for fewer small payments in strong disclosure states was physicians' reduced willingness to accept payments rather than the imposition of significant administrative costs on industry. We conduct additional analysis holding fixed the cost for pharmaceutical companies of disclosing data, which was possible because Massachusetts began releasing payment data online during our sample period. Differences-in-differences analyses and multiple regression yield similar estimates for each payment category: Mandatory disclosure reduced payments for speaking and for meals but increased payments for consulting activities. Significant disclosure aversion reducing conflicts of interest is consistent with the policy goals of mandatory disclosure, though the increased payments among those receiving large payments

may have been unintended.

4. Patients for Purchase: The Effects of Pharmaceutical Company Payments on Physician Prescribing and Patient Outcomes (E. Reinhart) This project links administrative Medicare data to industry-physician relationships cleaned from litigation settlements (a comprehensive dataset is available through the Affordable Care Act with 25 million payments including the date of payment) to examine the impact of disclosure laws and the impact of pharmaceutical company payments to doctors on prescribing, patient outcomes, and patient adherence. Physician Learning (E. Gentry, M. Gentry) examines whether industry-physician relationships affect physician learning. Malpractice Risk of Treatment Choices: Evaluating Legal Cases with CMS Microdata (E. Reinhart) uses Medicare data from 2006-2012 and Florida malpractice claims to build a predictive model of malpractice risk at the physician level and examine its interaction with industry-physician relationships.
5. The alienability of legal claims holds the promise of increasing access to justice and fostering development of law. While much theoretical work points to this possibility, no empirical work has investigated the claims, largely due to the rarity of trading in legal claims in modern systems of law. Can Markets Stimulate Rights? On the Alienability of Legal Claims (*RAND Journal of Economics*, 46(1), 23-65, 2015) develops a principal-agent framework where litigation funders provide expertise in reducing uncertainty in agents' disutility of production. The model leads to the counterintuitive prediction that litigation funders prefer cases with novel issues, and social surplus is positively correlated with legal uncertainty. Consistent with the model, court backlog, court expenditures, and a slowing in average time to completion are associated with third-party funding; cases with third-party funding receive more citations and are reversed less often than comparable cases without such arrangements. A Market for Justice: A First Empirical Look at Third Party Litigation Funding (*Journal of Business Law*, 15(3), 2013; D. Abrams) takes the first step toward empirically testing some of these theoretical claims using data from Australia. We find some evidence that third party funding corresponds to an increase in litigation and court caseloads. While third party funding appears to have effects on both the cases funded and the courts in jurisdictions where it is most heavily used, the overall welfare effects are still ambiguous.
6. Ideas Have Consequences: The Impact of Law and Economics on American Justice (*Quarterly Journal of Economics*, reject & resubmit; E. Ash, S. Naidu) provides a quantitative analysis of the effects of the law and economics movement on the U.S. judiciary using the available universe of opinions in U.S. Circuit Courts and 1 million District Court criminal sentencing decisions linked to judge identity. We estimate the effect of attendance in the controversial Manne economics training program that 40% of federal judges attended by 1990. To isolate the effect of judges from the types of cases they face, we exploit random assignment of judges to control for court- and case-level factors, an exogenous seating network from random panel composition to trace the spread of economic reasoning in law, and ordering of cases within Circuit to identify general economic ideas that move across legal topics. We use natural language processing methods to quantify the influence of economics in written judicial opinions. Descriptively, we find that judges who use law and economics language vote for and author conservative verdicts (as coded by Songer-Auburn) in economics cases and are more opposed to government regulation. After attending Henry Manne's economics training program, participating judges use more economics language and render conservative verdicts in economics cases, rule against regulatory agencies, particularly in labor and environmental cases, get cited more and increase dissents. These results are robust to a large set of judge biographical controls, and do not exist prior to Manne program attendance, suggesting a causal effect of economics training on judicial decisions. Further, Manne economics training is more predictive of these decisions than appointing political party. We further document a number of indirect channels of economics influence on the law beyond the direct effect on Manne program participants. Non-Manne judges exposed to Manne peers on previous cases increase their

use of economics language in subsequent opinions. Further, some economics concepts are portable across legal contexts: The Propagation of Economic Ideology: Peer Effects in Language Use in U.S. Appeals Courts (E. Reinhart) shows that “general-purpose” economics phrases such as “capital”, and “efficiency” move across legal topics within a judge. Economics reasoning diffused from regulatory domains into criminal law. Consistent with this, we show that law and economics influenced criminal decisions: Circuit Court judges that attend the Manne program are more likely to reject criminal appeals, and this effect spills over onto non-Manne judges serving on the same panel, results robust to judge biographical controls and absent prior to Manne program attendance. Moving to district courts, after attending Henry Manne’s economics training program, participating judges increase sentence lengths, and using variation in judicial discretion generated by U.S. v. Booker, sentencing gaps between Manne and non-Manne judges grew by 20% (10 months) after this ruling, which allowed more judicial sentencing discretion. Finally, Laissez Faire Racism: The Exacerbation of Racial Inequalities by Economics Training for Judges, A Natural Experiment (E. Reinhart) documents that Manne attendance is more predictive of racial and gender sentencing disparities than party of appointment, and Manne judges in both Circuit and District Courts render harsher immigration decisions, voting for enforcement of immigration regulation and longer sentences for illegal aliens.

7. Training Effective Altruism (*Journal of Development Economics*, revise & resubmit; S. Naseer, S. Mehmood) randomizes different schools of thought on cultivating pro-sociality and finds that training the utilitarian value of empathy elevates pro-sociality among high-stakes decision makers in Pakistan. One month after training, treated civil servants display 0.4-0.6 sigma greater altruism; two to six months after, orphanage visits and blood donations double. Field and lab results suggest improved theory of mind in strategic dilemmas: blood donations only increased when treated individuals were told their exact blood type was in need, and training improved cooperation, coordination, and guessing the decisions of others (Nagel 1995). Training also increased language of social cohesion in social media.
8. Training Policy-Makers in Econometrics (S. Naseer, S. Mehmood). Training deputy ministers in the school of thought associated with the credibility revolution increases demand for and responsiveness to causal evidence. We used a simplified Becker Degroot Marshak lottery to randomize policymakers into an econometrics training program. Treated individuals’ beliefs about the importance of quantitative evidence in policy making increases by 1.32 sigma, performance in national research methods and public policy exams improves by 0.5-0.8 sigma, and willingness-to-pay for commissioning RCTs using public funding increases by 300%, One year after the training, in their official duties, they are twice as likely to support policies for which there is RCT evidence.
9. Transmitting AI Training: Evidence from Policymakers in Pakistan (S. Mehmood; S. Naseer) How does AI training transmit among policymakers and impact the population? What are the unintended consequences of different types of AI training? We randomly assign a rigorous “AI for policy” workshop to deputy ministers in Pakistan and find that deputy ministers shift their attitudes towards AI and increase funding for digitization. Amid land record digitization efforts, treated ministers’ jurisdictions reduced delays in handling land disputes by 33%. Four months after the workshop, we cross-randomized ministers to receive AI fairness activism arguments that emphasizes the inescapability of algorithmic bias. AI fairness activism causes policymakers to report greater costs associated with AI in policymaking and decrease funding for digitization. Both interventions transmit from the deputy ministers to their subordinate staff and impact the population. Overall, our results underscore how ideas are readily transmissible within the State and highlight the inescapability of algorithmic bias in AI may have unintended policy consequences in developing countries, stymieing important policy reforms such as land record digitization and decreasing citizen satisfaction with the State.
10. Transmitting Rights: Effective Cooperation, Inter-Gender Contact, and Student Achievement (*JPE Micro*, reject & resubmit; S. Naseer, S. Mehmood) This paper provides experimental evidence of teacher-to-student

transmission of gender attitudes in Pakistan. We randomly show teachers a pro-women’s rights visual narrative. Treated teachers increase their own and students’ support for women’s rights, unbiasedness in gender IATs, and willingness to petition parliament for greater gender equality. Students improve their coordination and cooperation with the opposite gender in strategic dilemmas. Effects compound when teachers teach a gender-rights curriculum. Mathematics achievement also increases for classrooms assigned to form mixed-gender study groups, while no significant effects appear under same-sex study groups. Overall, we show gender attitudes are transmissible, inter-gender contact benefits learning, and non-cognitive skills like cooperation improve student outcomes.

11. Why Are Rights Revolutions Rare? (S. Naseer, S. Mehmood) We show that when traditional norms are challenged, norm disruptors pay a price. We experimentally foster more progressive gender attitudes among female teachers in Pakistan and show that these attitudes transmit to their students. Progressive gender attitudes, however, elevate stress hormone concentrations in blood plasma by 0.3 standard deviations and lead to a 0.35 standard deviation increase in domestic violence. Leveraging random variation in the fraction of teachers treated within a school, we find, however, that when additional teachers hold progressive attitudes—a moral bandwagoning effect—the costs of holding progressive gender attitudes are attenuated. Overall, our results suggest that deviation from traditional gender norms comes at a cost, but this cost diminishes as societal norms converge. In particular, domestic violence, more so than stress, stymies gender rights revolutions.
12. Role Models Matter for Covid Vaccinations and Conditional Cash Transfers Do Not: Impact on Vaccinations and Student Achievement (S. Naseer, S. Mehmood) The COVID-19 pandemic has taken an unprecedented hit on student learning across the globe. This paper provides experimental evidence that targeted messaging from role models can cause large reductions in vaccine hesitancy and improve student achievement. We randomly assign teachers to a menu of cash incentives, role models and celebrity messaging. One year later, teachers assigned to the role model treatment were 0.35 standard deviation units more likely to get vaccinated, had 0.5 standard deviation lower absenteeism, and had students with 0.15 standard deviation higher test scores. Our use of QR-code verified vaccine certificates provides a reliable and robust tracker of vaccination status. In the months following the treatment, we see effects of role model messaging grow on vaccination rates, teacher absenteeism, and student test scores. As vaccination becomes widespread, the effect of role model messaging converges for vaccination and absenteeism, but the impacts on student achievement persist. Behavioral data on theory-of-mind suggests the mechanism underlying the role model effect. Overall, our results suggest that an effective teacher vaccination campaign may attenuate learning losses among students due to COVID-19.
13. Homophily and Transmission of Behavioral Traits in Networks (P. Bhargava†, M. Sutter, C. Terrier) Social networks are a key factor of success in life, but they are also strongly segmented on gender, ethnicity, and other demographic characteristics (Jackson, 2010). We present novel evidence on an understudied source of homophily, namely behavioral traits. Behavioral traits are important determinants of life-time outcomes. While recent work has focused on how these traits are influenced by the family environment or how they can be affected by childhood interventions, little is still known about how these traits are associated to social networks. Based on unique data that we collected using incentivized experiments on more than 2,500 French high-school students, we find high levels of homophily across all ten behavioral traits that we study (including social, risk, competitive preferences, and aspirations). Notably, the extent of homophily depends on similarities in demographic characteristics, in particular with respect to gender. Furthermore, the larger the number of behavioral traits that students share, the higher the overall homophily. Then, using network econometrics, we show that the observed homophily is not only an outcome of endogenous network formation, but is also a result of friends influencing each others’ behavioral traits. Importantly, the transmission of traits is larger when students share demographic characteristics, such as gender, have been friends for longer or are friends

with more popular individuals.

14. Grit and Academic Resilience During Covid-19 (PLoS-ONE, revise & resubmit; S. Ertac, T. Evgeniou, A. Nadaf, X. Miao, E. Yilmaz†)

6.2 Distinguishing Sacred Values from Social Preferences: Theory, Evidence, and Relevance of Deontological Motivations

Economics and law are famously distinguished by consequentialism and deontological approaches to policy. This research asks the behavioral question and considers its implications. In the last few decades, there has been a gradual expansion of the domain of preferences considered by formal theory. There is the homo-oconomics view – that people are only motivated by material consequences of their decisions on their own payoffs. Then, models were expanded to incorporate fairness into economics, for example, caring about the consequences of their decisions for others and what people think of my intentions. We consider another slice, that is, can people be motivated simply out of duty—purely internal consequences? In general, consequentialism and deontological motivations are very hard to distinguish. What we propose is a method using revealed preference to identify non-consequentialist motivations and a model of deontological motivations as lexicographic (“first, do your duty”). An approximation of lexicographic preferences is a concave cost of deviating from duty. In economics it is standard to model costs as being convex. This is often plausible, reflecting the idea that small deviations from a bliss point are rather negligible, while larger deviations may have big consequences. However, there is one domain in which experimental data seem to suggest that costs are in fact concave—the domain of ideology and morality. Popular terms like the “What the Hell Effect” and “Slippery Slope” capture exactly the perception that individuals are perfectionist in this domain, and they do not distinguish much between small and large deviations from their bliss points. The cognitive cost of wrongdoing in a large scale is not much larger than in a small scale, hence the need for perfectionists to be careful when considering even a small deviation from the “right” principle. Whether the costs of deviating from ideological or moral bliss points is convex or concave has implications for how to integrate different cultures and religion, and how to address extremism (rightwing, Islamists, etc.) in particular. Different sanctioning structures (be it interpersonal, informal, or coming from an authority) affect the choice of individuals with non-majoritarian ideologies to integrate. In particular, being harsh on small cultural deviations may backfire and create subcultures. This framework forms the basis for studying more detailed questions of collision between cultures and it provides a large number of implications that can be tested both empirically and experimentally.

1. Social Preferences or Sacred Values? Theory and Evidence of Deontological Motivations (*Science Advances*, 8(19), eabb3925, 2022; M. Schonger) proposes an experimental method that can detect an individual’s deontological motivations by varying the probability of the decision-maker’s decision having consequences. It uses two states of the world, one where the decision has consequences and one where it has none. We show that a purely consequentialist decision-maker whose preferences satisfy first-order stochastic dominance will choose the decision that leads to the best consequences regardless of the probability of the consequential state. A purely deontological decision-maker is also invariant to the probability. However, a mixed consequentialist-deontological decision-maker’s choice changes with the probability. The direction of change gives insight into the location of the optimand for one’s duty. We provide a formal interpretation of major moral philosophies and a revealed preference method to detect deontological motivations and discuss the relevance of the theory and method for economics and law.
2. A Theory of Experiments: Invariance of Equilibrium to the Strategy Method of Elicitation and Implications for Social Preferences (*Journal of the Economic Science Association*, revise & resubmit; M. Schonger) develops implications of deontological motivations for economics methodology. Most papers that employ the strategy

method (SM) use many observations per subject to study responses to rare or off-equilibrium behavior that cannot be observed using direct elicitation (DE), but ignore that the strategic equivalence between SM and DE holds for the monetary payoff game but not the game participants actually play, which is in terms of utilities. To illustrate the severity of this issue, we formalize the mapping from the monetary payoff game to this actual game. A theorem provides necessary and sufficient conditions for strategic equivalence to apply. When the domain of preferences includes commonly-modeled motivations, such as intentions or disappointment aversion, or less-common ones, such as self-image or duty, strategic equivalence fails and thus the invariance to the method of elicitation does not apply. We use results from the past literature and our own experiments to investigate how well this theorem explains when results with SM and DE differ. We manipulate the salience of off-equilibrium considerations in our own experiments to demonstrate that SM and DE are not strategically equivalent, contrary to conventional wisdom. Three results emerge. First, not accounting for the bias in the estimation when decisions at one information set can influence the utility at another information set can render significant differences in decision-making. Second, the bias can be large and equivalent to some of the other causal effects being measured. Third, subtle interventions on salience can magnify these differences by a similar amount.

3. Who Cares? Measuring Attitude Strength in a Polarized Environment (*Journal of Politics*, under review; C. Cavaille and K. Van der Straeten) builds on research in social psychology, we propose a model of survey response in which individuals' policy preferences are characterized by two parameters: their attitude on an issue, and their attitude strength. Strong attitudes are behaviorally relevant and stable over time while weak attitudes are easily manipulated with only limited behavioral consequences. We assume that the psychological cost to individuals of not reporting their attitude depends positively on the issue strength. We derive predictions about how respondents will answer survey questions under two different survey techniques, Likert scales and Quadratic Voting (QV). The QV method gives respondents a fixed budget to "buy" votes in favor or against a set of issues. Because the price for each vote is quadratic, it becomes increasingly costly to acquire additional votes to express support or opposition to the same issue. We formally show that QV better measures preference strength. This, we argue, is especially true in a polarized two-party system where individuals with weak preferences are more likely to mechanically default to the party policy position. Results from a population-representative survey provides support for our argument.
4. Willingness to Say? Optimal Survey Design for Prediction (*American Economic Review: Insights*, under review; C. Cavaille, R. Das, K. Van der Straeten) Survey design often approximates a prediction problem: the goal is to select instruments that best predict an unobserved construct or future outcome. We demonstrate how advances in machine learning techniques can help choose among competing instruments. First, we randomly assign respondents to one of four survey instruments to predict a behavior defined by our validation strategy. Next, we assess the optimal instrument in two stages. A machine learning model first predicts the behavior using individual covariates and survey responses. Then, using doubly robust welfare maximization and prediction error from the first stage, we learn the optimal survey method and examine how it varies across education levels.
5. Quadratic Voting for Survey Research (QVSR): A Platform for Deploying QV in Survey Contexts (*AI Magazine*; NeurIPS 2021 (AI for Credible Elections); M. Bassetti, R. Das, G. Dias, A. Mortoni) Civic engagement is increasingly becoming digital as the ubiquity of computing in everyday life increases technologically mediated daily interactions. Governments have been instating various digitization efforts to harness these new facets of virtual life. What remains to be seen is if citizen political opinion, which can inform the inception and effectiveness of public policy, is being accurately captured. Civichase is an open-source online platform for deploying surveys that use decision-making mechanisms and analytical techniques at the forefront of civic

- engagement. Its focus is to aid researchers implement interactive surveys that involve the application of Quadratic Voting Survey for Research (QVSR) at scale, and with an intuitive, delightful user experience. Furthermore, in this paper, we explore the optimal survey design for prediction, which seeks to understand if QVSR vs other elicitation methods is optimal in predicting political behaviour. Finally, we investigate Personal AI, which is an emerging domain, and its relevance to modelling individual political preferences, and how this could be implemented using an existing participatory governance and machine learning framework.
6. A Decision-Theoretic Approach to Understanding Survey Response: Likert vs. Quadratic Voting for Attitudinal Research (*University of Chicago Law Review Online*, 22(2019); C. Cavaille, K. Van der Straeten) "Likert scales" are the most standard and widespread instrument in survey research when measuring public opinion on political and economic issues. In this simple approach, respondents are given the opportunity to voice their agreement or disagreement on a set of issues by placing their attitudes on a scale that runs from "strongly disagree" to "strongly agree". One assumption commonly made by social scientists using such scales is that they provide faithful - if noisy - measures of respondents' views. We challenge this assumption, highlighting several reasons why respondents may be expected to systematically exaggerate their views in political surveys using Likert scales. We propose a simple decision-theoretic model of survey answers to discuss whether Quadratic Voting might overcome these pathologies. We provide conditions under which one might expect Quadratic Voting to outperform Likert scales.
 7. Non-Confrontational Extremists (*European Economic Review*, revise & resubmit; M. Michaeli, D. Spiro) develops another implication of deontological motivations for economics methodology. We present a critique of a standard assumption in economics—that the cost of deviating from one's bliss point is convex—with fundamental implications for decision making in social and political settings and for the empirical predictions of theoretical models in these domains. We show that the concavity of ideological and moral costs holds in (1) naturally generated data, (2) where stakes are high, (3) when decisions are made by experts, and (4) when the consequences have the effect of silencing people with opinions that are far from the consensus, because they are the ones for whom it is most difficult to stand on their ground. In a high-stakes field setting (U.S. Courts of Appeals), we examine which individuals, on an ideological scale, conform most to the opinions of others. Legal precedents are set by ideologically diverse and randomly composed panels of judges. Using exogenous predictors of ideology and rich voting data we show that ideological disagreements drive dissents against the panel's decision, but ideologically extreme judges are caving in: they are the least likely to dissent and their voting records are the least correlated with their predicted ideology. Meanwhile, moderately ideological judges are dissenting the most despite evidence that they are more often determining the opinion. Our theoretical analysis shows that these findings are most consistent with a model of decision making in the presence of peer pressure with a concave cost of deviating from one's ideological convictions—perfectionism.
 8. Attitudes as Assets (S. Naseer, S. Mehmood, A. Seror) Attitudinal survey data has been a foundation for researchers, policymakers, and democracies. Yet preference falsification, the act of misrepresenting a preference under public pressure, is common. We micro-found attitudinal data collection into two components: the utility of expressing the underlying preference weighed against the price of expressing that attitude. We randomly expose fundamentalist or progressive speech to rural teachers in Pakistan and show that fundamentalist speech does not affect teachers' preferences although it makes fundamentalist attitudes less costly. The progressive speech makes teachers' preferences more liberal. We validate our results with petitions to the parliament.
 9. Modularizing the code across experiments led to oTree: An Open Source Platform for Online, Lab, and Field Experiments (*Journal of Behavioral and Experimental Finance*, 9(1), 88-97, 2016; M. Schonger, C. Wickens).

oTree is an open-source and online software for implementing interactive experiments in the laboratory, online, the field or combinations thereof. oTree does not require installation of software on subjects' devices; it can run on any device that has a web browser, be that a desktop computer, a tablet or a smartphone. Deployment can be internet-based without a shared local network, or local network-based even without internet access. The programming language is Python, a popular, open-source programming language. www.otree.org provides the source code, a library of standard game templates and demo games which can be played by anyone. As of 2016, oTree has been translated into French, German, Italian, Japanese, Spanish, and Russian. It is used in Computer Science, Economics, Engineering, Mathematics, Physics, Political Science, Psychology, and Sociology for classroom, lab, MTurk, and field experiments without internet. It is used in Australia (Melbourne, Newcastle, Perth, Queensland), Austria (Innsbruck, Vienna), Belgium (Leige), Canada (Guelph, Toronto), China (Beijing), Czech (Prague), Finland (Aalto), France (CReA Defense, Lille, Montpellier, Nice, Toulouse), Germany (GfK Marketing Research, Mannheim, Munich), Hungary (Academy of Sciences), Italy (Bologna, European University Institute), Japan (Tokyo), Kenya (Nairobi), Korea (Seoul), Netherlands (Amsterdam, Maastricht, Nijmegen, Tilburg, United Nations University, Utrecht), Norway (Norwegian School of Economics), Russia (RANEPA), South Africa (Pretoria), Spain (Madrid, Malaga, Valencia, Zaragoza), Sweden (Gothenburg, Stockholm), Switzerland (ETH, Geneva, Lausanne, Zurich), U.K. (Cambridge, Lancaster, Oxford), and U.S. (Boston College, Colby, Columbia, Iowa State, Michigan State, Northwestern, NYU, Ohio State, Princeton, Treasury Department, UCSC, UCSD, University of Chicago, Vanderbilt, Stanford, Yale).

7 Behavioral Judging

By justice, we think of equal treatment before the law and equality based on recognition of difference. We can imagine a set of covariates X that should lead to the same predictions, or predictability, of outcomes $Y = f(X) + \varepsilon$. X should improve the predictions. And there's a set of W 's that should not ($y \perp W, \text{var}(\varepsilon) \perp W$). We tend to think of the X 's as consequences of actions ($a \rightarrow X, a \not\rightarrow W$), based of the control principle and merit principle: individuals should be responsible only for events that are under their control. The W 's can be race, gender, expressions of one's identity, and factors outside of one's control - football, weather, lunchtime, outcomes of the preceding case.

Judges who have self-concept of being just, if they misperceive what is a fair or random sequence may actively negatively autocorrelate trying to do what they think is right. Judges who think of themselves as moral and have lexicographical preferences—first, do your duty—and approximate these lexicographical preferences with a concave cost of deviating from your duty may intentionally cave in when the benefits of deviation is too high. Judges who see people like themselves, randomly, excessively, may start to distinguish themselves; dissimulation in the management of self-identity. Judges may assign longer sentences to individuals who threaten their ego. Judges can also indirectly bias when they say they're impartial but they are actually partial, leaving bias unexamined.

7.1 Impossibility of Objective Judgments? Priming, Gambler's Fallacy, Mood, Voice, and Peer Effects in U.S. Courts

1. Decision-Making Under the Gambler's Fallacy: Evidence From Asylum Courts, Loan Officers, and Baseball Umpires (*Quarterly Journal of Economics*, 2016, 131(3): 1181-1241; T. Moskowitz, K. Shue) finds consistent evidence of negative autocorrelation in decision-making that is unrelated to the merits of the cases considered in three separate high-stakes field settings: refugee asylum court decisions, loan application reviews, and major league baseball umpire pitch calls. The evidence is most consistent with the law of small numbers and the gambler's fallacy – people underestimating the likelihood of sequential streaks occurring by chance – leading to negatively autocorrelated decisions that result in errors. The negative autocorrelation is stronger among

more moderate and less experienced decision-makers, following longer streaks of decisions in one direction, when the current and previous cases share similar characteristics or occur close in time, and when decision-makers face weaker incentives for accuracy. Other explanations for negatively autocorrelated decisions such as quotas, learning, or preferences to treat all parties fairly, are less consistent with the evidence, though we cannot completely rule out sequential contrast effects as an alternative explanation.

2. Recent work in cognitive science provides overwhelming evidence for a link between emotion and moral judgment. Mood and the Malleability of Moral Reasoning: The Impact of Irrelevant Factors on Judicial Decision Making (*Journal of Public Economics*, under review; M. Loecher) detects intra-judge variation spanning three decades in 1.5 million judicial decisions driven by factors unrelated to case merits. U.S. immigration judges grant an additional 1.4% of asylum petitions—and U.S. district judges assign 0.6% fewer prison sentences and 5% longer probation sentences—on the day after their city’s NFL team won, relative to days after the team lost. Bad weather has the opposite effect of a team win. Unrepresented parties in asylum bear the brunt of NFL effects. The effect on district judges only appears for judges born in the same state as the current state of residence, providing clean evidence of extraneous influences on judge decision-making as opposed to lawyer or applicant behavior. Moving beyond OLS, we utilize models from machine learning to estimate the sentence length relative to the sentencing guideline. We find that while several appropriate features predict sentence length, such as details of the crime committed, other features seemingly unrelated, including daily temperature, sport game scores, and location of trial, are predictive as well. The predictive power of the unrelated events is derived from the permutation based variable importance score in random forests. We address recent criticism of the reliability of these scores with double residualization.
3. Clash of Norms: Judicial Leniency on Defendant Birthdays (*Journal of Economic Behavior and Organization*, revise & resubmit; A. Philippe) documents judicial leniency on defendant birthdays across 5 million decisions. French sentences are 1% fewer and 3% shorter. U.S. federal sentences are 33% shorter in the day component of sentences (the month component remains unaffected). New Orleans sentences are 15% shorter overall. No leniency appears on the days before or after a defendant’s birthday. Federal judges using deterrence language in opinions, are unaffected, isolating the judicial as opposed to defendant channel. The effect is doubled when judge and defendant share the same race. Our courtroom setting rules out many models of social preferences with reciprocity motives.
4. Gender Attitudes in the Judiciary: Evidence from U.S. Circuit Courts (*American Economic Journal: Applied Economics*, forthcoming; E. Ash, A. Ornaghi) Do gender attitudes influence interactions with female judges in U.S. Circuit Courts? In this paper, we propose a novel judge-specific measure of gender attitudes based on use of genderstereotyped language in the judge’s authored opinions. Exploiting quasi-random assignment of judges to cases and conditioning on judges’ characteristics, we validate the measure showing that slanted judges vote more conservatively in gender-related cases. Slant influences interactions with female colleagues: slanted judges are more likely to reverse lower-court decisions if the lower-court judge is a woman than a man, are less likely to assign opinions to female judges, and cite fewer female-authored opinions.
5. Implicit egotism—in particular, unconscious associations that individuals have with others who share their names or first initials—is a mainstay of modern psychology textbooks. Using unique data on 48,988 randomly assigned defendants from 1988-1999, The Judicial Superego: Implicit Egoism, Internalized Racism, and Prejudice in Three Million Sentencing Decisions (S. Hajdini, E. Reinhart) finds that judges assign 8% longer sentences to defendants whose first initials match their own. Name letter effects amplify when the first and second letter of the name match, when the entire name matches, when the name letter is rare, and appear for roughly all judges in the sample. The effects are larger for black defendants classified as “Negro” rather

than “Black”. The first initial effect and the difference by racial label replicates for the last name. The effect is consistent with self-image motivations to create social distance from negatively-valenced targets perceived to be associated with the self.

6. The emphasis on “fit” as a hiring criterion has raised the spectrum of a new form of subtle discrimination. Under complete markets, correlations between malleable characteristics and outcomes should not persist. Yet using data on 1,901 U.S. Supreme Court oral arguments between 1998 and 2012, Covering: Mutable Characteristics and Perceptions of Voice in the U.S. Supreme Court (Y. Halberstam, A. Yu) documents that voice-based snap judgments based on lawyers’ identical introductory sentences, “Mr. Chief Justice, (and) may it please the Court?”, predict court outcomes. The connection between vocal characteristics and court outcomes is specific only to perceptions of masculinity and not other characteristics, even when judgment is based on less than three seconds of exposure to a lawyer’s speech sample. Consistent with employers mistakenly favoring lawyers with masculine voices, perceived masculinity is negatively correlated with winning and the negative correlation is larger in more masculine-sounding industries. The first lawyer to speak is the main driver. Among these petitioners, males below median in masculinity are 7 percentage points more likely to win in the Supreme Court. Female lawyers are also coached to be more masculine and women’s perceived femininity predict court outcomes. Republicans, more than Democrats, vote for more feminine-sounding females, while Democrats, but not Republicans, vote for less masculine-sounding men. A de-biasing experiment finds that information reduces 40% of the correlation between perceived masculinity and perceived win, and incentives reduce another 20%. A model shows how the information treatment identifies statistical discrimination and the incentives treatment identifies taste. Perceived masculinity explains additional variance relative to and is orthogonal to the best random forest prediction model of Supreme Court votes. In a separate paper written for a general scientific audience, Perceived Masculinity Predicts U.S. Supreme Court Outcomes (*PLoS-ONE*, 11(10), e0164324; Y. Halberstam, A. Yu) documents the correlation between vocal characteristics and court outcomes. While this study does not aim to establish any causal connections, our findings suggest that vocal characteristics may be relevant in even as solemn a setting as the Supreme Court of the United States. Previous studies suggest a significant role of language in the court room, yet none has identified a definitive correlation between vocal characteristics and court outcomes.
7. Using data from 1946–2014, Attorney Voice and the U.S. Supreme Court (in *Law as Data*, Santa Fe Institute Press, forthcoming, ed. M. Livermore and D. Rockmore; Y. Halberstam, M. Kumar, A. Yu) shows that audio features of lawyers’ introductory statements and lawyers’ facial attributes improve the performance of the best prediction models of Supreme Court outcomes. We infer face attributes using the MIT-CBCL human-labeled face database and infer voice attributes using a 15-year sample of human-labeled Supreme Court advocate voices. We find that image features improved prediction of case outcomes from 63.8% to 65.6%, audio features improved prediction of case outcomes from 66.8% to 68.8%, image and audio features together improved prediction of case outcomes from 66.9% to 67.7%, and the weights on lawyer traits are approximately half the weight of the most important feature from the models without image or audio features. Predictions of Justice votes with image and audio features however remained more similar relative to their baselines. We interpret this difference to suggest that human biases are more relevant in close cases.
8. Electoral Cycles Among U.S. Courts of Appeals Judges (*Journal of Law and Economics*, 60(3), 479-496, 2017; C. Berdejo) finds evidence consistent with experimental studies that document the contexts and characteristics making individuals more susceptible to priming. Just before U.S. Presidential elections, judges on the U.S. Courts of Appeals double the rate at which they dissent and vote along partisan lines. Increases are accentuated for judges with less experience and in ideologically polarized environments. During periods of national reconciliation—wartime, for example—judges suppress dissents, again, especially by judges with less

experience and in ideologically polarized environments. We show the dissent rate increases gradually from 6% to nearly 12% in the quarter before an election and returns immediately to 6% after the election. If highly experienced professionals making common law precedent can be politically primed, it raises questions about the perceived impartiality of the judiciary.

9. Why do presidential elections affect judges without electoral incentives? In randomly composed 3-judge panels, U.S. Circuit judges' dissents increase for ten months preceding elections and peak at the nominating conventions. Reversals of lower courts and political partisanship in precedents by unified panels double. Priming Ideology: Why Presidential Elections Affect U.S. Judges (*Science Advances*, under review) develops a theoretical model showing that the salience of partisan identities can explain this pattern. Using novel administrative linkages between judges' states and case progression from trial, circuit, to (potentially) the Supreme Court, this article documents polarizing effects that vary by intensity of elections across states, within judge, and over the electoral season. Exploiting monthly campaign ads in judges' states of residence, dissent elevation is higher in electorally pivotal states, but declines precipitously in non-pivotal states after the primary season. Topic of dissents, placebo dates from earlier case milestones, replication in concurrences (disagreement about reasoning), replication with U.S. Senate elections, and judge heterogeneity support a transient priming mechanism. Elections explain 23% of all dissents.
10. Many legal decisions, such as whether to set bail or release on parole, are made as part of a sequence of similar but independent decisions. Does the serial position of a case within a sequence influence the decision? Previous research in non-legal domains mostly suggests that cases appearing later in the sequence are likely to be judged more favorably than cases appearing early in the sequence. To check for the effect of serial position on legal decisions, Best to be Last: Serial Position Effects in Legal Decisions in the Field and in the Lab (*Journal of Applied Psychology*, forthcoming; O. Plonsky, Y. Feldman, T. Steiner, L. Nitzer) takes a dual approach. First, we analyzed a real-world dataset of refugee asylum court cases over a period of 33 years (N = 386,109). The results show that the higher the serial position of an asylum request in a given day, the more likely it is to be granted, in line with sequential decisions research in other domains. To complement these findings, we ran three controlled experiments (N = 1,872) in which laypeople faced sequences of legal cases and were asked to make hypothetical choices. The results of all three experiments show the same pattern: decisions become more favorable later in the sequence. Our dual analysis of real-world observational data and carefully designed controlled experiments thus suggests that, from the point of view of the affected individual, it is best to be last.
11. Time of Day and Decision Heuristics: Evidence from Asylum Judges finds judges grant asylum more before lunch and less after. They also grant more asylum towards the end of day. These time of day effects are independent from sequence effects.
12. Judicial Leniency Grows with Age shows that judges become more lenient in three different settings—they grant more asylum, they assign shorter sentences in federal district courts, and they are more likely to reverse a lower court decision in criminal appeals in the circuit courts.
13. The Relativity of Racial Perception: Color Contrast Effects in Refugee Courts (E. Reinhart) finds evidence of sequential contrast effects along the racial dimension in asylum courts.
14. Peer effects is further explored in Social Contagion and Political Ideology: Evidence from Repeated Random Exposure in the U.S. Courts of Appeals (E. Reinhart) and polarization in The Relationality of Judgement: Social Dynamics of Opinion-Formation in U.S. Courts of Appeals (E. Reinhart). Using repeated random assignment to teams, the effect of being minority finds instead of assimilation, we see dis-assimilation; the

effect of being majority finds instead of integration, we see radicalization; and the effect of uniformity finds instead of conformity, we see egotism.

15. Supreme Court Vacancies and Discretionary Opinion Writing in Federal Circuit Courts (W. Lu) studies the behavior of Circuit Court judges during Supreme Court vacancies. Judges who were candidates of nomination converge to mimic other judges in dissents, concurrences, and voting for the US government.
16. When Matching Markets Unravel: Theory and Evidence from Federal Judicial Clerkships (Y. He, T. Yamashita) studies the judge-clerk match, a market plagued by unraveling. Evidence from a unique dataset on match and production shows that (1) agents on either side have similar preferences over those on the other side, (2) the matching game for judges is close to zero-sum, (3) this fierce competition among judges explains the unraveling in this market. We develop a theoretical model investigating how homogeneity of preferences (and competition) affects unraveling in matching markets. We show that a static mechanism, as proposed in many previous reforms, is impossible to solve the problem of unraveling in a market with a high degree of homogeneity. By contrast, a dynamic mechanism that takes advantage of judges' repeated participation in the market over time is proven promising. Based on our findings, we propose a new market design for the judge-clerk match.
17. Rituals (*Nature Human Behavior*, forthcoming; S. Mehmood, A. Seror) We study the effect of religious practice on judicial behavior in the context of Ramadan fasting, in the second largest Muslim majority country in world. For identification, we rely on exogenous variation in the length of daily fasting due to the interaction between the rotating Islamic calendar and a district court's geographical latitude. Using unique case-level microdata on criminal cases spanning the entire history of Pakistan and random assignment of cases across Muslim and non-Muslim judges, we find that in contrast to the literature on physiological deprivation making judges more severe, more intense fasting increases acquittals. Religion appears to drive our results, since we find no effect of Ramadan on the judicial behavior of non-Muslim judges and in cases involving violations of Islamic Law. We replicate these findings in India with a much larger sample.
18. Measuring Gender and Religious Bias in the Indian Judiciary (E. Ash, S. Asher, A. Bhowmick, T. Devi, C. Goessmann, P. Novosad, B. Siddiqi) We study judicial in-group bias in Indian criminal courts, collecting data on over 80 million legal case records from 2010–2018. We exploit quasi-random assignment of judges and changes in judge cohorts to examine whether defendant outcomes are affected by being assigned to a judge with a similar religious or gender identity. We estimate tight zero effects of in-group bias. The upper end of our 95% confidence interval rejects effect sizes that are one-fifth of those in most of the prior literature.
19. Who is in justice? Caste, religion and gender in the courts of Bihar over a decade (*Economics and Political Weekly*, under review; S. Bhupatiraju, S. Joshi, P. Neis) Bihar is widely regarded as one of India's poorest and most divided states. It has also been the site of many social movements that have left indelible marks on the state's politics and identity. Little is currently known about how structural inequalities have affected the functioning of formal systems of justice in the state. We use a novel dataset of more than 1 million cases filed at the Patna high court between 2009—2019 together with a variety of supplementary data to analyze the role of religion, caste and gender in the high court of Bihar. We find that the courts are not representative of the Bihari population. Muslims, women and scheduled castes are consistently under-represented. The practice of using "caste neutral" names is on the rise. Though there is little evidence of "matching" between either judges and petitioners or between judges and filing advocates on the basis of names, we do find evidence that petitioners and their advocates match on the basis of identity such as the use of "caste neutral" names. These results suggest that the social movements which disrupted existing social structures in the past may

have inadvertently created new social categories that reinforce networks and inequalities in the formal justice system.

20. The Promise of Machine Learning for the Courts of India (*National Law School of India Review*, forthcoming; reprinted as Government Analytics Using Machine Learning (*Handbook of Measurement*; S. Bhupatiraju, S. Jankin, G. Kim, M. Kupi, M. Ramos-Maqueda) Artificial Intelligence (AI) and machine learning (ML) – adaptive computer programs that can perform functions typically associated with the human mind – offer new opportunities for improving the courts of India. We argue that the integration of machine learning tools with newly available legal data offers a mechanism to identify biases in judicial behavior and propose real-time corrections for these behaviors, thereby streamlining the system and reducing backlog. We describe some of the methods for constructing a robust pipeline to scrape, clean and prepare this data in India for analysis. We conclude with directions—from automated causal impact analyses of judicial rulings to randomized controlled trials that carefully estimate the direct and indirect impacts of the adoption of these algorithms, as well as their cost-effectiveness.

7.2 Judicial Analytics and Behavioral Foundations of Polarization

1. Ambiguity aversion has recently been proposed as a behavioral foundation for polarization. Ambiguity aversion has also been used to explain a wide range of phenomena in law and policy: incomplete contracts, stock market volatility, abstention from voting, why prosecutors offer and defendants accept harsh plea bargains—and why individuals update priors after receiving identical information in an opposing manner. Is Ambiguity Aversion a Preference? Ambiguity Aversion Without Asymmetric Information (*Journal of Behavioral and Experimental Economics*, revise & resubmit; M. Schonger) presents evidence problematizing the experimental basis for ambiguity aversion. Ambiguity aversion is the interpretation of the experimental finding known as the Ellsberg paradox that most subjects violate probabilistic sophistication: They prefer betting on events whose probabilities are known (objective) to betting on events whose probabilities are unknown to them (subjective). However in typical experiments these unknown probabilities are known and often determined by the experimenter. Thus the typical Ellsberg experiment is a situation of asymmetric information. People may try to avoid situations where they are the less informed party in an asymmetric situation setting. Indeed doing so is often normatively appropriate. Thus avoidance of situations of informational asymmetry is a potential confound in typical Ellsberg experiments. Paying to avoid information asymmetry in an Ellsberg experiment would constitute the misapplication of a heuristic to the unfamiliar experimental situation. To eliminate this confound, this paper proposes a new source of ambiguity: participant generated ambiguity. Instead of the experimenter filling an Ellsberg urn, the opaque Ellsberg urn is filled by the other subjects in a laboratory session. We find that eliminating asymmetric information while leaving ambiguity in place, makes subjects more than willing to choose the ambiguous bet rather than the objective one even when choosing the objective bet is costless.
2. Testing Axiomatizations of Ambiguity Aversion (*Theory and Decision*, revise & resubmit; M. Schonger) empirically interrogates existing ambiguity aversion theories. The study of the normative and positive theory of choice under uncertainty has made major advances through thought experiments often referred to as paradoxes: the St. Petersburg paradox, the Allais paradox, the Ellsberg paradox, and the Rabin paradox. Machina proposes a new thought experiment which posits a choice between two acts that have three outcomes. As in the Ellsberg paradox there are three events, but while the Ellsberg paradox has two (monetary) outcomes in Machina there are three. Machina shows that four prominent theories of ambiguity aversion predict indifference between the acts. Introspection, however, suggests that many people might very well strictly prefer one act over the other. This paper makes four contributions: first, to our knowledge, it is the first to experi-

mentally implement the Machina thought experiment. Second, we employ a novel method to simultaneously elicit the certainty equivalent of an embedded lottery. Third, our results—across three experiments—indicate non-indifference, which rejects earlier theories of ambiguity aversion, but is consistent with a newer one, which we apply to explain our results. Fourth, we show that independence is a sufficient condition for indifference in the Machina paradox, and thereby explains why so many models predict indifference.

3. Motivated Reasoning in the Field: Polarization of Precedent, Prose, and Policy in U.S. Circuit Courts, 1930-2013 (W. Lu) documents motivated reasoning among U.S. Circuit Court judges. We employ a supervised learning approach to measure partisan influences on prose (writing style), precedent (citations to previous cases), and policy (dissenting votes). We find persistent but low partisanship of language overall, with the notable exception of Civil Rights and First Amendment jurisprudence. Citations display a significant level of partisanship and increase over time. Voting along party lines (dissenting against judges from opposing party) and strategic retirement (retiring while one's own party controls the presidency) have also increased. These measures are predictive of ideological direction of future voting. Finally, we show that motivated reasoning grows with judicial experience, but not age, and is more pronounced for Republican appointees.
4. A case study in motivated reasoning is observed in that less than 1% of U.S. Federal judges report political motivations for retirement and resignation. However using two centuries of data, The Disavowal of Decisionism in American Law: Political Motivation in the Judiciary (*Journal of Law and Courts*, reject & resubmit; E. Reinhart) shows that 13% of retirements and 36% of resignations follow political cycles. When the President comes from a *different* political party as judge's party of appointment, U.S. Courts of Appeals judges are *less likely to retire* in each of the three quarters *before* a Presidential election. In contrast, judges are *more likely to resign* in each of the four quarters *after* a Presidential election when the President comes from the judge's party of appointment. Politically motivated exits have increased significantly in recent years to constitute one-fifth of retirements since 1975. We are able to uncover these patterns by analyzing the data at the quarter-to-election level, while prior research has relied on self-reports or yearly analysis. We expand on the implications of this finding in A Judge Retires. Just How Political Is That Decision? (*The New York Times*, April 14, 2022; E. Reinhart).
5. What Kind of Judge is Brett Kavanaugh? A Quantitative Analysis (*Cardozo Law Review*, 2018; E. Ash) reports the results of a series of data analyses of how recent Supreme Court nominee Brett Kavanaugh compares to other potential Supreme Court nominees and current Supreme Court Justices in his judging style. The analyses reveal a number of ways in which Judge Kavanaugh differs systematically from his colleagues. First, Kavanaugh dissents and is dissented against along partisan lines. More than other Judges and Justices, Kavanaugh dissents at a higher rate during the lead-up to elections, suggesting that he feels personally invested in national politics. Far more often than his colleagues, he justifies his decisions with conservative doctrines, including politicized precedents that tend to be favored by Republican-appointed judges, the original Articles of the Constitution, and the language of economics and free markets. These findings demonstrate the usefulness of quantitative analysis in the evaluation of judicial nominees. Kavanaugh is radically conservative. Here's the data to prove it (*Washington Post*; Jul 10, 2018, E. Ash) presents a summary.
6. Federal courts are a mainstay of the justice system in the United States. In What Matters: Agreement Between U.S. Courts of Appeals Judges (*Journal of Machine Learning Research (W&CP)*, 2016; X. Cui, L. Shang, J. Zheng), we analyze 387,898 cases from U.S. Courts of Appeals, where judges are randomly assigned to panels of three. We predict which judge dissents against co-panelists and analyze the dominant features that predict such dissent with a particular attention to the biographical features that judges share. Random forest achieves the best classification. Dissent is roughly half-driven by case features and half-driven by personal features.

7. Precedent vs. Politics? Case Similarity Predicts Supreme Court Decisions Better Than Ideology uses the universe of U.S. Circuit Court cases appealed to the Supreme Court since 1946. We show that case similarity among Circuit Court opinions achieves better prediction accuracy of Supreme Court decisions relative to the current best prediction model, which is based on ideology of judges and trends of how they vote. Relative to the benchmark prediction accuracy of 59%, textual measures of case similarity achieve prediction accuracy of 64%. We interpret this improvement to suggest that precedent matters more than politics alone. Combining case similarity with ideological features further improves accuracy to 72%, suggesting that ideology affects interpretation of precedent. We also offer our model available as a web app.
8. Affirm or Reverse? Using Machine Learning To Help Judges Write Opinions predicts higher court reversals of lower court decisions. Every year more than 300,000 civil and criminal cases are heard in the district courts all over the U.S. Less than 5% of these cases are appealed and heard in circuit courts. For most of the cases, the circuit court either affirms the decision of the district court or reverses it. Out of the cases heard in circuit courts, only about 2% are heard in the Supreme Court. The Supreme Court again can again either affirm or reverse the circuit court decision. We build a model to predict the higher court decision using the lower court's opinion using a digital corpora of circuit court cases to 1891 and district court cases to 1923. Comparing a wide variety of classification, dimensionality reduction, and oversampling techniques, we are able to achieve an accuracy of 79% in circuit courts and 68% accuracy in the Supreme Court.
9. Predicting Punitiveness and Sentencing Disparities from Judicial Corpora uses over 1 million sentencing decisions linked to judge identity and the digital corpora of U.S. district court opinions. We show that judges' writings can predict average harshness and racial and sex disparities in sentencing decisions. We document significant reductions in mean square error relative to a naive prediction (the mean of the training data) on the test dataset by approximately 24 percent in predicting punitiveness.
10. Implicit Bias in Supreme Court Speech: Inferences of Gender Attitudes from Vocal Patterns Predict Judicial Decisions uses 15 years of Supreme Court oral arguments. We show that vocal intonation of gendered words (e.g., actor vs. actress) classify vocal intonations of neutral words into stereotypically male (e.g., logical, ability, think) and female (e.g., looking, cook, goodwill), surprisingly suggesting the relevance of people's perceptions of gender being revealed in how people speak. Furthermore, the vocal intonations of judges' speaking these words are predictive of their decisions. Justice Stephen Breyer can be predicted with 73% accuracy while Justice Antonin Scalia can be predicted with 58% accuracy. These results complement other work indicating that perceptions of gender improve predictions of Supreme Court outcomes and continue to play a role in a manner more complex and nuanced than conventionally perceived. Analysis of Vocal Implicit Bias in SCOTUS Decisions Through Predictive Modeling (*Proceedings of Experimental Linguistics*, 2018; R. Vunikili, H. Ochani, D. Jaiswal, R. Deshmukh, E. Ash) leveraged machine learning techniques to detect bias in the judicial context by examining the oral arguments. The adverse implications due to the presence of implicit bias in judiciary decisions could have far-reaching consequences. This study aims to check if the vocal intonations of the Justices and lawyers at the Supreme Court of the United States could act as an indicator for predicting the case outcome.
11. Dialects of Ideology examines whether speech variation beyond word choice, that is, fluctuations in the way one speaks holding the words fixed is predictive of ideology in the U.S. Supreme Court. We use lawyers' campaign donations as a commonly-used measure of political ideology. We find that audio significantly improves prediction accuracy of ideology relative to using the text alone. AUC increases from 0.55 to 0.61, even in a setting as solemn as the Supreme Court.
12. Investigating Variation in English Vowel-to-Vowel Coarticulation in a Longitudinal Phonetic Corpus (*Proceed-*

ings of the 18th International Congress of Phonetic Sciences, 2015; C. Abrego-Collier, J. Phillips, B. Pillion, A. Yu) investigates the nature of individual variation in speech, particularly the mechanism underlying such variability, is increasingly important, especially for research on sound change, since such investigations might help explain why sound change happens at all and, conversely, why sound change is so rarely actuated even though the phonetic preconditions are always present in speech. The present study contributes to the literature on inter- and intra-speaker variation in coarticulation, a major precursor to sound change, by focusing on the degree of coarticulation stressed vowels have on neighboring unstressed vowels using recordings from a longitudinal phonetic corpus of oral arguments before the Supreme Court of the United States. Significant inter-speaker variation in height coarticulation, both anticipatory and carryover, is observed, while no evidence for systematic inter-speaker variability in backness coarticulation is found. There is also no evidence for intra-speaker variation in coarticulation over the course of 205 days. Mimicry: Phonetic Accommodation Predicts U.S. Supreme Court Votes (A. Yu) digitizes speech patterns in US Supreme Court oral arguments and shows that lawyers converge to justices are more likely to win their vote, and that justices who converge to each other during oral arguments are more likely to vote together.

13. Making Information Actionable: Evidence from a Nationwide Experiment in Kenyan Courts (M. Chemin, M. Ramos-Maqueda). Courts in developing countries face numerous constraints that prevent them from providing efficient and fair justice to citizens, or a strong institutional environment conducive to investment and growth. We ask: can low-cost, information-based interventions, using data regularly captured by administrative systems, help judicial officers overcome common incentive and behavioral constraints, and in so doing improve court performance?
14. Courts and Informality Across Countries (J. Lee, P. Neis) Informality and rms: Many developing countries share two stylized facts. First, much or most of their workforce is informally employed. This is especially important since informality can have negative impacts for both the workers themselves (for instance see Gindling and Newhouse (2014)) and for the state, which collects less taxes and social security contributions. Second, many of these countries have a relatively independent judiciary which are characterized by high, and often increasing, backlogs. Slow courts can have a direct impact on people's lives and rms' decisions. Courts' impacts on firms can happen through multiple potential channels and observed for instance as underinvestment or missallocation due to a worse contract enforcement. Despite much work on links between weak institutions and informality and despite a deep empirical and theoretical understanding of dierent topics such as institutions in general, informality, firms' behavior or the judiciary itself, little is known about the link between the judiciary and (worker) informality. In this paper, we shed light on the relationship between court efficiency and the share of informal labor across the world. We look at aggregate measures in five countries (Chile, Croatia, India, Kenya and Peru) and strengthen our analysis with state-level data from India. This allows us to compare countries in dierent states of economic development, with different historical backgrounds and different judicial systems and to search for common links between informality and judicial efficiency across them.
15. Impact of Free Legal Search on Rule of Law: Evidence from Indian Kanoon (S. Bhupatiraju, R. Das, S. Joshi, P. Neis) Access to legal information is limited in many parts of the world. Can digital platforms offering free legal search reduce market-level constraints on economic development? We estimate the impact of Indian Kanoon, a free legal search engine, using a generalized difference-in-differences empirical strategy. We find that the staggered rollout of Kanoon was associated with a 1-2% increased likelihood of case resolutions and doubling of the number of appeals, which are also less likely to be dismissed by the courts. It also affected the finances of firms with positive impacts on assets and negative impacts on audit fees and bad debts.

16. Improving Legal Training: the Impact of Social-Emotional Learning and Class Monitoring on Judicial Performance (B. Silveira, M. Ramos-Maqueda). Despite the importance of high-stake judicial decisions on litigants' wellbeing and economic development, there is little evidence on how to improve judicial performance. In this paper, we use a randomized controlled trial to evaluate the impact of teacher monitoring and social-emotional exercises on the performance of judges and prosecutors in Peru's Judicial Academy. We test the impact of our interventions on both soft and hard skills that aim to improve not only educational attainment and social-emotional skills, but also contribute to fairer and more efficient judicial decision-making.
17. Information Provision and Court Performance: Experimental Evidence from Chile (P. Carrillo, B. Silveira, M. Ramos-Maqueda). Previous studies have shown that behavioral nudges can be a cost-effective tool to influence changes in people's actions. In this study, we aim to test whether nudging court managers through informing them on how their court performs in absolute and relative terms can improve court productivity. Moreover, we test if there is any difference if the information about the court performance is given in contrast and relation to self past performance or if the information is relative to other courts' performance.

8 Law and Legitimacy

According to communitarian philosophers, the idea of ego or self-concept can be traced to Enlightenment and Romantic ideals of the self—Enlightenment ideals of self-knowledge and self-mastery, presuming that each person has an interior space, and memory of oneself, and Romantic ideals of self-expression and authenticity, being true to oneself. Recognizing that there are different sources of the self, according to philosophers, constituted a moral revolution, in which projects of personal identity came to be important beyond economic self-interest. Questions of identity became personally significant; refusals of acceptance and respect, deeply challenging. We can see projects of personal identity being mobilized in progressive (the transformation of gender and sexual identities and claims to equal rights) and conservative politics (the claims to national and communal identities defended by populist movements). We can see projects of personal identity in self-esteem and authenticity being taught as values in school (“find your passion”). We can see projects of personal identity and self-esteem—violations of self-esteem—in contemporary discussions of micro-aggression, trigger warnings, and privilege disparities. According to these philosophers, recognizing that everyone has their own way of being human facilitates respect for individuals, but also for different cultures.

Whether the law is legitimate can be related to how individuals are treated, the identity of the lawmaker, and one's reference points. It can be studied in contextualized field experiments and observed in failed policy changes. Most research on judicial decision-making examine when extraneous factors influence their decisions; inverting the question, evidence when legal factors should affect their decisions, but they do not, raise questions of snap or pre-determined judgement.

8.1 Deter or Spur: British Executions During World War I

Deter or Spur? British Executions During World War I is a book-length version of *The Deterrent Effect of the Death Penalty? Evidence from British Commutations During World War I*. During World War I, the British military condemned over 3,000 soldiers to death, but only executed 12% of them; the others received commuted sentences. Many historians believe that the military command confirmed or commuted sentences for reasons unrelated to the circumstances of a particular case and that the application of the death penalty was essentially a random, “pitiless lottery.” Over two death sentences a day precluded careful consideration to execute or commute by the Commander-in-Chief, who was responsible for the final decision.

In addition, we have the names of all the soldiers—those who received the death sentence and those who

deserted—and we can see, to a first approximation, if they were Irish or British. The Irish were a large subordinated minority that declared independence right after the war. Like minorities elsewhere, they were sentenced to death at a much higher rate than the general population. But conditional on the death sentence, they were equally likely to be executed. Holding punishment constant allows the potential identification for a channel beyond deterrence, namely legitimacy, to explain compliance to the law.

Many psychologists, sociologists, political scientists, and legal theorists have all emphasized the importance of legitimacy in courts, organizations, and nation states. But the existing quantitative evidence has been limited to surveys asking different communities whether they think the lawgiver is legitimate and correlating these perceptions with rates of crime, which can be correlated for unobserved reasons. In particular, they can be correlated if unequal punishment leads to higher incarceration rates and lower perceptions of legitimacy or if economic inequality drives crime rates and perceptions. This identification of the potential deterrent and delegitimizing effects of executions hinges on whether executions were random.

Using a dataset on all capital cases during World War I, I find that the data are consistent with an essentially random process. Using this result, I exploit variation in commutations and executions within military units to identify the deterrent effect of executions, with deterrence measured by the elapsed time within a unit between the resolution of a death sentence (i.e., a commutation or execution) and subsequent absences within that unit. Absences are measured via handwritten trial records and “wanted” lists prepared by British military police units searching for deserters and preserved in war diaries and police gazettes. I find some limited evidence that executing deserters deterred absences, while executing Irish soldiers, regardless of the crime, spurred absences, particularly Irish absences. I present a model where perceived legitimacy of authority affects why people obey the law.

The differences between the situation I study and contemporary criminal justice scenarios are vast, so a more nuanced understanding of the differences is required in order to draw policy lessons from the WWI experience. Because the British experience provides an extremely low-bar test for the death penalty, finding a deterrence effect in the context of WWI would certainly not be a strong argument, leaving aside moral issues, that the death penalty is good policy. However, a negative result showing no deterrent effect might have more policy salience since if we ever expected to find an effect, it would be in the WWI context: executions took place almost immediately—in a manner purposefully designed to maximize their deterrent effect—and death sentences were given out very frequently and quite arbitrarily for military desertion. We would still expect that on the margin more executions should deter absences and if we find this not to be the case, it would suggest that the threat of future death for crimes is not as strong a disincentive as we might imagine. Despite these differences, this study offers some insights potentially capable of greater generalization. The granularity and richness of the data begets questions that are sometimes ignored in standard crime rate studies.

8.2 Incarceration And Its Disseminations

1. Incarceration And Its Disseminations: COVID-19 Pandemic Lessons From Chicago’s Cook County Jail (*Health Affairs*, 2020; E. Reinhart) argues that jails and prisons are major sites of novel coronavirus (SARS-CoV-2) infection. Many jurisdictions in the United States have therefore accelerated the release of low-risk offenders. Early release, however, does not address how arrest and pretrial detention practices may be contributing to disease spread. Using data from Cook County Jail—one of the largest known nodes of SARS-CoV-2 spread in the United States—in Chicago, Illinois, we analyzed the relationship between jailing practices and community infections at the ZIP code level. We found that jail–community cycling was a significant predictor of cases of coronavirus disease 2019 (COVID-19), accounting for 55 percent of the variance in case rates across ZIP codes in Chicago and 37 percent of the variance in all of Illinois. Jail–community cycling far exceeds race, poverty, public transit use, and population density as a predictor of variance. The data suggest that cycling people through Cook County Jail alone is associated with 15.7 percent of all documented COVID-19 cases

in Illinois and 15.9 percent of all documented cases in Chicago as of April 19, 2020. Our findings support arguments for reduced reliance on incarceration and for related justice reforms both as emergency measures during the present pandemic and as sustained structural changes vital for future pandemic preparedness and public health.

2. Epidemiological Consequences of Jail Cycling in Marginalized Communities: Criminal Punishment and Structural Racism during Covid-19 (*Proceedings of the National Academy of Sciences*, 118 (21), e2026577118, 2021; E. Reinhart) argues that jails and prisons remain leading sites of Covid-19 outbreaks. Mass incarceration poses ongoing health risks for communities. We investigate whether short-term jailing of individuals prior to release may drive Covid-19 spread in communities. We find that the cycling of individuals through Cook County Jail in March 2020 alone can account for 13% of all Covid-19 cases and 21% of racial Covid-19 disparities in Chicago as of early August. We conclude that detention for alleged offenses that can be safely managed without incarceration is likely harming public safety and driving racial Covid-19 disparities. These findings reinforce consensus among public health experts that large-scale decarceration should be implemented to protect incarcerated people, mitigate Covid-19 spread and racial disparities, and improve future pandemic preparedness.
3. Effects of Jail Decarceration and Anti-Contagion Policies on Covid-19 Spread in the United States (*JAMA Network Open*, 4(9), e2123405, 2021; E. Reinhart) provides the first empirical analysis of the epidemiological effects of jail decarceration (alongside 10 major anti-contagion policies) on Covid-19 spread in US counties. We find that reducing jail populations by 80%—a reduction that would bring the US closer to average incarceration rates among peer nations and that is feasible through alternate management of non-violent alleged crimes—would have reduced daily Covid-19 growth rates by 2%. Because growth rates compound, such a reduction would have prevented millions of Covid-19 cases. Decarceration is a key policy intervention for public health.

8.3 Revealed Preference Indifference: Legitimacy, Law, and Recognition-Respect

1. Machine Learning and Rule of Law (in *Law as Data*, Santa Fe Institute Press, forthcoming, ed. M. Livermore and D. Rockmore) argues that predictive judicial analytics holds the promise of increasing the fairness of law. Much empirical work observes inconsistencies in judicial behavior. By predicting judicial decisions—with more or less accuracy depending on judicial attributes or case characteristics—machine learning offers an approach to detecting when judges most likely to allow extra legal biases to influence their decision making. In particular, low predictive accuracy may identify cases of judicial “indifference,” where case characteristics (interacting with judicial attributes) do not strongly dispose a judge in favor of one or another outcome. In such cases, biases may hold greater sway, implicating the fairness of the legal system.
2. Early Predictability of Asylum Court Decisions (*Proceedings of the ACM Conference on AI and the Law*, 2019; M. Dunn, L. Sagun, H. Sirin) presents evidence of judges ignoring information. In the United States, foreign nationals who fear persecution in their home country can apply for asylum under the Refugee Act of 1980. Over the past decade, legal scholarship has uncovered significant disparities in asylum adjudication by judge, by region of the United States in which the application is filed, and by the applicant’s nationality. These disparities raise concerns about whether applicants are receiving equal treatment under the law. Using machine learning to predict judges’ decisions, we document another concern that may violate our notions of justice: we are able to predict the final outcome of a case with 80% accuracy at the time the case opens using only information on the identity of the judge handling the case and the applicant’s nationality. Moreover, there is significant variation in the degree of predictability of judges at the time the case is assigned to a judge.

We show that highly predictable judges tend to hold fewer hearing sessions before making their decision, which raises the possibility that early predictability is due to judges deciding based on snap or predetermined judgments rather than taking into account the specifics of each case. Early prediction of a case with 80% accuracy could assist asylum seekers in their applications.

3. *The Legal Reproduction of Racism: Determinants of Sentencing Disparities* (E. Reinhart) introduces individual-level prosecutorial (as well as individual-level judicial) data. I observe screening (including declination decisions), charging, trial, and sentencing outcomes in a unique dataset vertically linked from the time of arrest. I leverage the random assignment of cases to prosecutors and judges and ask whether judge-based or prosecutor-based disparities interact in a courtroom setting – i.e., does the race of the prosecutor influence the extent of judge-based disparities. Two results emerge. First, cases assigned to white judges and black prosecutors receive shorter sentences. Second, black-white sentencing disparities reverse for cases assigned to black judges and black prosecutors. These results are consistent with narratives of racial hierarchy in the law whereby black-white disparities are rendered and reproduced.
4. *How Prosecutors Exacerbate Racial Disparities* (E. Reinhart) links the universe of individuals in a district attorney’s office over a decade with many stages of random assignment. Three facts emerge. First, racial disparities in criminal sentencing magnify once discretion in prosecutorial screening is taken into account. 50% of arrestees are screened out by prosecutors, who disproportionately favor whites. Taking this screening gap into account greatly magnifies real sentencing disparities. Second, race of screener prosecutors matters significantly. White prosecutors screen in fewer arrestees and these defendants get assigned shorter sentence lengths. Black prosecutors screen in far more black arrestees. Third, racial interactions in the courtroom affect sentencing levels and disparities. When defendants are assigned to black trial prosecutors and white judges, they receive substantially lighter sentences. Black trial prosecutors and black judges eliminate or reverse the sentencing gap, such that longer sentences are rendered for white defendants. This eliminates the theoretical possibility that black-white sentencing disparities reflect unobserved heterogeneity beyond race.
5. *Self-Corrosion of Law: Effects of Criminal Justice Exposure on Perceptions of Law’s Legitimacy* (A. Philippe, E. Reinhart) studies the long-run effects of criminal justice exposure using the random assignment of prosecutors with different leniency on trust in the law and voting behavior.
6. The appointment of Sonia Sotomayor to the Supreme Court in 2009 was criticized as sacrificing merit on the altar of identity politics. According to critics, Sotomayor was simply “not that smart.” For some conservative critics, her selection illustrated the costs of affirmative action policies, in that this particular choice was going to produce a lower quality Supreme Court. For liberal critics, many were concerned that the President, by selecting Sotomayor, was squandering an opportunity to appoint an intellectual counterweight to conservative Justices like Antonin Scalia, Samuel Alito, and John Roberts. Using a set of basic measures of judicial merit, such as publication and citation rates for the years 2004 to 2006, when Sotomayor was on the Court of Appeals for the Second Circuit, ‘Not that Smart’: Sonia Sotomayor and the Construction of Merit (*Emory Law Journal*, 61(4), 2012, with G. Charles and M. Gulati) compares her performance to that of her colleagues on the federal appeals courts. Sotomayor matches up well.
7. *Tastes for Desert and Placation: A Reference Point-Dependent Model of Social Preferences* (*Research in Experimental Economics*, Experimental Economics and Culture, Volume 20, 205-226, 2018; Bingley, UK: Emerald; ed. A. Gunthorsdottir and D. A. Norton) proposes a model of behavior in social interactions where individuals maximize a three-term utility function: a conventional consumption utility term and two “social” terms that capture social preference. One social term is a taste for desert, which is maximized when the individual believes the other person is getting what they deserve. The second social term measures the

target individuals' anger or gratitude from the interaction which is determined by a value function derived from prospect theory. After introducing the model and generating a series of comparative statics results and derived predictions, I report the results of a series of quasi-field experiments on social preferences. I discuss how the model explains several paradoxes of empirical moral philosophy that are less explicable by current economic models of social preference focusing on outcomes and intentions.

8. Legal theorists have suggested that literature stimulates empathy and affects moral judgment and decision-making. Law and Literature: Theory and Evidence on Empathy and Guile (*Review of Law and Economics*, 15(1), 2018) presents a model to formalize the potential effects of empathy on third parties. Empathy is modeled as having two components—sympathy (the decision-maker's reference point about what the third party deserves) and emotional theory of mind (anticipating the emotions of another in reaction to certain actions). I study the causal effect with a data entry experiment. Workers enter text whose content is randomized to relate to empathy, guile, or a control. Workers then take the Reading the Mind in the Eyes Test (RMET) and participate in a simple economic game. On average, workers exposed to empathy become less deceptive towards third parties. The result is stronger when workers are nearly indifferent. These results are robust to a variety of controls and model specifications.
9. The emotions that someone expresses has consequences for how that person is treated. The Strategic Display of Facial Expressions (*Leadership Quarterly*, revise & resubmit; A. Hopfensitz, J. Van Der Ven, B. Van Leeuwen) studies whether people display emotions strategically. In two laboratory experiments, participants play task delegation games in which managers assign a task to one of two workers. When assigning the task, managers see pictures of the workers and we vary whether getting the task is desirable or not. We find that workers strategically adapt their emotional expressions to the incentives they face, and that it indeed pays off to do so. Yet, workers do not exploit the full potential of the strategic display of emotions.
10. Reforms can be rejected if perceived to be illegitimate. Interim Report on a Preschool Intervention Program in Kenya (resting paper; P. Glewwe, M. Kremer, S. Moulin) evaluates an educational program that professionalized an informal educational system. Teacher training, classroom materials, and incentives for teacher attendance was provided to fifty preschools randomly selected from one hundred preschools in rural Kenya. Teachers were eligible for bonuses of up to 85% of pre-program salary depending on their attendance. Headmasters acted as monitors and distributed funds. In practice, headmasters typically paid the entire bonus to teachers regardless of attendance, which tended to crowd out parental contributions to teacher salary. Teacher training significantly reduced the number of minutes spent on the blackboard. The point estimates suggest that the program improved teacher attitude, energy, effort, control, and organization. The program increased progression to grade one by the end of three years; however, it also significantly decreased written test scores after two years.
11. Legitimizing Policy (*American Political Science Review*, under review; M. Michaeli, D. Spiro) In many settings of political bargaining over policy, agents care not only about getting their will but also about having others approve the chosen policy thus giving it more weight. What is the effect on the bargaining outcome when agents care about such legitimacy of the policy? We study this question theoretically and empirically. We show that the median-voter theorem holds in groups that are either very cohesive or have extreme ideological disagreement. However, in groups with intermediate ideological disagreement, non-median agents can—and do—affect the policy. This is since, on the individual level, ideological disagreement with the median has a non-monotonic effect on the policy. We test our model in a natural experimental setting—U.S. appeals courts—where causal identification is based on random assignment of judges into judicial panels of three, and where judges care about legitimacy of the policy. The empirical tests corroborate our theoretical predictions.

12. Contract Enforcement in a Stateless Economy (S. Mehmood) How do markets flourish in the absence of a formal State authority enforcing contracts. To study contract enforcement without the State, we focus on a large informal market: illegal gambling. We present evidence that even in the absence of legal enforcement authority, personal relationships, and violence, more than 70% of gamblers fulfill their contractual obligations in this informal economy. We provide experimental evidence that even in impersonal and informal markets, reputation plays a key role in the honoring of contractual obligations. Extensions of the payment deadline also increases contract enforcement. Overall, we provide data and causal evidence that contract enforcement is possible in impersonal markets and without formal legal enforcement. Illegal gambling thrives on the same principles of reputation and credit constraints that sustain modern legal markets.

9 Demography of Ideas

The final section explores the transmission and persuasion process of novel moral theories, whose incommensurability can lead to group conflict. On a thick vein, I study the macro and demographic forces and laws that aid or hinder discrimination in legal institutions, markets, and public policy.

1. Developing countries with highly unequal income distributions, such as Brazil or South Africa, face an uphill battle in reducing inequality. Educated workers in these countries have a much lower birthrate than uneducated workers. Assuming children of educated workers are more likely to become educated, this fertility differential increases the proportion of unskilled workers, reducing their wages, and thus their opportunity cost of having children, creating a vicious cycle. Income Distribution Dynamics with Endogenous Fertility (*American Economic Review*, 89(2), 155-160, May 1999; NBER Working Paper No. w7530; M. Kremer) and Income Distribution Dynamics with Endogenous Fertility (*Journal of Economic Growth*, 7(3), 227-258, 2002; NBER Working Paper No. w7530; M. Kremer) present a model incorporating this effect, which generates multiple steady-state levels of inequality, suggesting that in some circumstances, temporarily increasing access to educational opportunities could permanently reduce inequality. Empirical evidence suggests that the fertility differential between the educated and uneducated is greater in less equal countries, consistent with the model.
2. A number of countries have begun implementing tax incentives designed to reverse the decline in fertility. Whether such incentives are effective or equitable remains an open question. During the early 20th century, France initiated an unusual tax policy to promote fertility and marriage: household income was divided by family size to obtain a final tax bracket. Similar policies whose fertility benefit increases with income are being implemented today. Using hand-collected archival data from aggregate tax returns and three natural experiments, Can Countries Reverse Fertility Decline? Evidence from France's Marriage and Baby Bonuses, 1929-1981 (*International Tax and Public Finance*, 118(3), 252-271, 2011) finds mixed evidence that these tax incentives affect fertility and marriage.
3. For those facing infertility, using assisted reproductive technology to have genetically related children is a very expensive proposition. In particular, to produce a live birth through in vitro fertilization (IVF) will cost an individual (on average) between \$66,667 and \$114,286 in the U.S. If forced to pay these prices out of pocket, many would be unable to afford this technology. Given, this reality, a number of states have attempted to improve access to this technology through state-level insurance mandates that cover IVF. Several scholars, however, have worried that increasing access in this way will cause a diminution in adoptions and have argued against enactment of these mandates for that reason. In Trading Off Reproductive Technology and Adoption: Do IVF Subsidies Decrease Adoption Rates? (*Minnesota Law Review*, 95(2), 2010, I. G. Cohen), we push against that conclusion on two fronts. First, we interrogate the normative premises of the argument

and expose its contestable implicit assumptions about how the state should balance the interests of existing children waiting for adoption and those seeking access to reproductive technology in order to have genetically related children. Second, we investigate the unexamined empirical question behind the conclusion: does state subsidization of reproductive technologies through insurance mandates actually reduce adoption; that is, is there a trade-off between helping individuals conceive and helping children waiting to be adopted? We call the claim that there is such an effect the “substitution theory.” Using the differential timing of introduction of state-level insurance mandates relating to In Vitro Fertilization (IVF) in some states and differences in the forms these mandates take, we employ several different econometric techniques (differences-in-differences ordinary least square, two-stage least square) to examine the effect of these mandates on IVF utilization and adoption. Contrary to the assumption of the substitution theory, we find no strong evidence that state support of IVF through these mandates crowds out either domestic or international adoption.

4. On a practical level, concepts of family formation and views on sexual conduct and freedom may be fundamentally divided on religious conflict lines—some scholars contend that divisions between Western and Muslim countries are vast when it comes to attitudes towards divorce, abortion, gender equality, and gay rights. Different views of family can be decisive for the support for inter-group transfers within a nation and could decide on support for, for example, universal welfare goods, public provision of pensions and health-care. There may also be related conflict risks (e.g., between groups where each group views the other as immoral). Different views on family and religious denomination could also be important for willingness to trust one another and support for public goods. These issues are particularly important for countries that are experiencing a potential polarization of attitudes (such as parts of Europe, where relatively large proportions of the population are highly liberal, while several important subpopulations and growing migrant groups have traditional family views).

9.1 Genealogy of Ideology

1. Building towards the demography of attitudinal change is a long-term research goal. The incommensurability of novel moral theories can lead to group conflict, which I have studied using data in objective and subjective knowledge in science and law. Thomas Kuhn’s *Structure of Scientific Revolutions* (1962) proposes a theory of knowledge—science undergoes periodic paradigm shifts—that has been interpreted to mean that there is no ‘truth’ in science. To test this theory, *How Does Science Progress? A Statistical Approach to Postmodern Theories of Knowledge* constructs “citation trees,” where each node is a paper connected to all the papers it cites and the papers that cite it. Just as in evolutionary biology where a species tree has branches that get sparse or discontinued during periods of mass extinctions, I look for similar events in a citation tree. Using ISI data, I construct geology, linguistics, and literary criticism trees dating from 2001 back to 1945, 1956, and 1975 respectively. I also use the NBER Patent Citation database from 1975-1999. Markov clustering algorithms marking death of “bushes” indicate linguistics and geology are qualitatively different from literary criticism, where statistical “extinctions” occur often, contrary to the constructionist claim that all forms of discourse accumulate similarly. Paradigm shifts incommensurate enough to cause mass extinctions do not appear to have occurred in linguistics, geology, or patent citations but literary criticism may have had a paradigm shift in 1990.
2. *The Genealogy of Ideology: Identifying Persuasive Memes and Predicting Agreement in U.S. Circuit Courts* (*Proceedings of the ACM Conference on AI and the Law*, 2017; A. Parthasarathy, S. Verma) identifies memes based on the likelihood of legal phrases appearing along an edge of the citation tree but not appearing on a disconnected part of the tree. We then predicted how two judges in a particular panel align on their voting,

based on the historical vote alignment of that judge with other judges, historical citation of each other, and historical use of shared legal phrases and memes.

10 AI and Rule of Law

1. Judicial Analytics and the Great Transformation of American Law (*Journal of Artificial Intelligence and the Law*, 27(1), 15-42, 2019) presents an overview of ongoing machine learning analysis of 12 terabytes of judicial data. Predictive judicial analytics holds the promise of increasing efficiency and fairness of law. Judicial analytics can assess extra-legal factors that influence decisions. Behavioral anomalies in judicial decision-making offer an intuitive understanding of feature relevance, which can then be used for de-biasing the law. A conceptual distinction between inter-judge disparities in predictions and inter-judge disparities in prediction accuracy suggests another normatively relevant criterion with regards to fairness. Predictive analytics can also be used in the first step of causal inference, where the features employed in the first step are exogenous to the case. Machine learning thus offers an approach to assess bias in the law and evaluate theories about the potential consequences of legal change.
2. Algorithms as Prosecutors: Lowering Rearrest Rates Without Disparate Impacts and Identifying Defendant Characteristics Noisy to Human Decision-Makers investigates how machine learning might bring clarity to human decisions made during the criminal justice process. Our data comes from all cases at the New Orleans District Attorney's office for the years 1988-1999. We exploit random assignment of prosecutors, prosecutorial discretion, and heterogeneity across prosecutors in charge rates to compare prediction models to judicial decision makers. Our model of defendant rearrest, trained using defendant and offense characteristics, selects higher-risk individuals to prosecute than its human counterparts did. In particular: given a set charge rate, our model would reduce rearrest rates by five to nine percentage points. This model could have several important policy implications: it might identify defendant characteristics that are particularly 'noisy' to prosecutors; it could suggest ways of alleviating criminal caseloads without increasing crime rates; and it might provide important insights into how a prosecutor's background relates to the quality and nature of their charging decisions.
3. In Can Machine Learning Help Predict the Outcome of Asylum Adjudications? (*Proceedings of the ACM Conference on AI and the Law*, 2017; J. Eigel), we analyzed 492,903 asylum hearings from 336 different hearing 2 locations, rendered by 441 unique judges over a thirty-two year period from 1981-2013. We define the problem of asylum adjudication prediction as a binary 4 classification task, and using the random forest method, we predict twenty-seven years of refugee decisions. Using only data available up to the decision date, our model correctly classifies 82 percent of all refugee cases by 2013. Our empirical analysis suggests that decision makers exhibit a fair degree of autocorrelation in their rulings, and extraneous factors such as, news and the local weather may be impacting the fate of an asylum seeker. Surprisingly, granting asylum is predominantly driven by trend features and judicial characteristics—features that may seem unfair—and roughly one third-driven by case information, news events, and court information. Ignoring information is related to motivated cognition (the interpretation of information in a polarized manner) and may be related to perfectionism. If one's actions are unlikely to change as a result of information acquisition, why update one's beliefs? In Endogenous Information Acquisition: Wikileaks State Department Cables Predict Asylum Decisions Almost Perfectly relative to the previous best prediction model of asylum court decisions of 80%, we show that refugee asylum outcomes are nearly perfectly predicted (with 98% accuracy) after incorporating US diplomatic communications captured in WikiLeak cables. We interpret this to suggest that international current events interact with domestic immigration decisions in a meaningful way. Successfully predicting these

decisions may be used to counsel refugees on their application’s chance of success.

4. Machine Prediction of Appeal Success in U.S. Asylum Courts (C. Andrus, D. Godevais, G. Ng) This paper asks whether the decisions of the appeals boards of U.S. Asylum Courts can be predicted using machine learning tools applied to information on the lower-court decisions. We use a new data set of 830,000 asylum appeals for the years 1985 through 2013. We show that the decisions of asylum appeals can be predicted with 80% accuracy and 0.85 AUC. Comparable performance is obtained using only decisions in previous years as training data. Important predictors include the nationality of the asylee and the identity of the lower court judge. Our model suggests that the individuals who do not appeal have a very low predicted success rate.
5. Recent work in natural language processing represents language objects (words and documents) as dense vectors that encode the relations between those objects. Mapping the Geometry of Law using Document Embeddings (*Science Advances*, reject & resubmit; S. Bhupatiraju, K. Venkataramanan) explores the vectorization of legal beliefs, with the goal of understanding judicial reasoning and the causal impacts of law. We illustrate the usefulness of these vectors in three ways. First, we show that they recover intuitive institutional connections between judges. Second, we show the vectors can be used as features in a decision prediction task. Third, we show that they can be used to measure implicit bias by judges toward women and racial minorities. Case Vectors: Spatial Representations of the Law Using Document Embeddings (in *Law as Data*, Santa Fe Institute Press, ed. M. Livermore and D. Rockmore, forthcoming; E. Ash) show that these vectors encode information that distinguishes courts, time, and legal topics. The vectors do not reveal spatial distinctions in terms of political party or law school attended, but they do highlight generational differences across judges. We conclude the paper by outlining a range of promising future applications of these methods.
6. Automated Fact-Value Distinction in Court Opinions (*European Journal of Law and Economics*, 1-17, 2020; E. Ash, Y. Cao) studies the problem of automated classification of fact statements and value statements in written judicial decisions. We compare a range of methods and demonstrate that the linguistic features of sentences and paragraphs can be used to successfully classify them along this dimension. The Wordscores method by Laver et al. (2003) performs best in held out data. In an application, we show that the value segments of opinions are more informative than fact segments of the ideological direction of U.S. Circuit Court opinions.
7. What modes of moral reasoning do judges employ? Automated Classification of Modes of Moral Reasoning in Judicial Decisions (*Computational Legal Studies*, 2018; E. Ash, N. Mainali, L. Meier) constructs a linear SVM classifier for moral reasoning mode trained on applied ethics articles written by consequentialists and deontologists. The model can classify a paragraph of text in held out data with over 90 percent accuracy. We then apply this classifier to a corpus of circuit court opinions. We show that the use of consequentialist reasoning has increased over time. We report rankings of relative use of reasoning modes by legal topic, by judge, and by judge law school.
8. Identifying Policy Levers: Automatic Case Classification and Validation using Citations builds inputs for end-to-end machine learning estimates of the causal impacts of law, we consider the problem of automatic case classification. We consider the quasi-supervised multi-class problem using as training set the Chicago Judges Project (a hand-coded dataset of thousands of cases in over 20 politically salient topics). Our model achieves 84% correct classification using only the opinion’s text. We show that, among citations used at least three times, 78% fall within the same cluster/predicted category.

10.1 Algorithms and Society

1. A Better Way to Onboard AI (*Harvard Business Review*, 2020; B. Babic, T. Evgeniou, A. Fayard) discusses a gradual approach for implementing AI for enhancing decision making, based on its unique abilities to continuously learn and evolve – much like a “living organism” does. The approach is sensitive to three key features of effective AI: cooperation, trust, and transparency. It can help organizations navigate common new pitfalls, such as algorithmic aversion, behavioral biases as well as potential discrimination and unfairness. This idea is fleshed out in a full-length article with behavioral economics foundations in Incremental AI (*Asian Journal of Law and Economics*, forthcoming).
2. What are the consequences of intermediating moral responsibility through complex organizations or transactions? Intermediated Social Preferences: Altruism in an Algorithmic Era (in *Advances in the Economics of Religion*, forthcoming, Palgrave, ed. J. P. Carvalho, S. Iyer, J. Rubin) examines individual decision-making when choices are known to be obfuscated under randomization. It reports the results of a data entry experiment in an online labor market. Individuals enter data, grade another individual’s work, and decide to split a bonus. However, before they report their decision, they are randomized into settings with different degrees of intermediation. The key finding is that less generosity results when graders are told the split might be implemented by a new procurement algorithm. Those whose decisions are averaged or randomly selected among a set of graders are more generous. These findings relate to “the great transformation” whereby moral mentalities are shaped by modes of (a)social interaction.
3. The emergence of online labor markets makes it far easier to use individual human raters to evaluate materials for data collection and analysis in the social sciences. Designing Incentives for Inexpert Human Raters (*Proceedings of the ACM Conference on Computer Supported Cooperative Work*, 2011, with A. Shaw and J. Horton) reports the results of an experiment—conducted in an online labor market—that measured the effectiveness of a collection of social and financial incentive schemes for motivating workers to conduct a qualitative, content analysis task. Overall, workers performed better than chance, but results varied considerably depending on task difficulty. We find that treatment conditions, which asked workers to prospectively think about the responses of their peers—when combined with financial incentives—produced more accurate performance. Other treatments generally had weak effects on quality. Workers in India performed significantly worse than U.S. workers, regardless of treatment group.
4. In some online labor markets, workers are paid by the task, choose what tasks to work on, and have little or no interaction with their (usually anonymous) buyer/employer. These markets look like true spot markets for tasks rather than markets for employment. Despite appearances, Are Online Labor Markets Spot Markets for Tasks? A Field Experiment on the Behavioral Response to Wages Cuts (*Information Systems Research*, 27(2), 403-423; J. Horton) find via a field experiment that workers act more like parties to an employment contract: workers quickly form wage reference points and react negatively to proposed wage cuts by quitting. However, they can be mollified with “reasonable” justifications for why wages are being cut, highlighting the importance of fairness considerations in their decision making. We find some evidence that “unreasonable” justifications for wage cuts reduce subsequent work quality. We also find that not explicitly presenting the worker with a decision about continuing to work eliminates “quits,” with no apparent reduction in work quality. One interpretation for this finding is that workers have a strong expectation that they are party to a quasi-employment relationship where terms are not changed, and the default behavior is to continue working.
5. Confusing Average and Marginal Tax Rates: Experimental Evidence examines if people react to labor market schedules in a manner that suggests confusion between average and marginal tax rates. I present individuals with identical payment schedules for data entry of a series of paragraphs, making it arguably harder for

individuals to be confused by average or marginal payment schemes. If people “iron”, then individuals presented with payment schedules displaying average payment rates should do more work than individuals presented with payment schedules displaying marginal payment rates. My first experiment demonstrates this. A natural follow-up question is whether individuals still “iron” when they are not presented with either average or marginal payment schedules. This may more accurately reflect an actual real-world setting where schedules are often hidden and difficult to compute. My second experiment finds that when workers are not primed with per paragraph payments, they enter significantly more paragraphs than when they are shown either average or marginal payment schedules.

6. *The Economics of Crowdsourcing: A Theory of Disaggregated Labor Markets* examines what protects anonymous individuals from appropriation in disaggregated labor markets. A new kind of economic organization is emerging: the information-processing disaggregated labor market. In these online markets, which are organized by for-profit firms acting as labor market intermediaries, workers are freelancers who perform tasks for requesters for either hourly rates or piece rates, sometimes with incentives for quality or speed. Somewhat ironically, these very new labor markets most resemble the simplest models of labor markets. The sociological and psychological aspects of traditional work relationships are largely absent: work is proposed on take-it-or-leave-it terms, and workers accept or reject offers based only on the onerousness of the work and the pay. There are no compensating differentials or benefits, no unions, no career concerns, and so on. What kinds of contractual mechanisms prevent the hold-up problem that would otherwise cause the market to unravel? Does the fixed price vs. cost plus nature of transactions in different disaggregated labor markets explain the contractual mechanisms that are actually observed? This paper presents some descriptive facts and a simple model illustrating the role of market design for these disaggregated labor markets.