CONTAGIOUS DISCRIMINATION: WHY RACE-BLIND JUSTICE IN COURTROOMS IS NOT ENOUGH*

YIJIA LU† AND MURAT C. MUNGAN‡

Existing theories of discrimination focus on discrimination at the individual, societal, or organizational level. We highlight the importance of the dynamic relationship between different institutional actors who jointly give rise to discriminatory outcomes. This reveals that discrimination can be contagious: discrimination that would otherwise not occur at one institutional level can arise due to discrimination at another institutional level. Therefore, discrimination by some actors can trickle down and create disparate impacts in processes that would otherwise be free of discrimination. This phenomenon, which we term "contagious discrimination," can manifest itself in various settings such as employment and commercial trade negotiations. To illustrate its importance, we focus on the dynamic relationship between plea-bargaining—a critical part of the criminal justice system—and policing. We demonstrate that a defendant who anticipates racially discriminatory arrest by the police may accept a less favorable plea-bargain from the prosecutor in comparison to a similarly situated defendant who does not expect such discrimination by the police. This occurs even when the prosecutor is completely race-blind, highlighting how discrimination by law enforcement can lead to discriminatory outcomes at the plea-bargaining stage, thereby illustrating that discrimination can be contagious. We discuss the implications for policies focused on defendants' perceptions, plea-bargain reform, and legal representation.

---

† Assistant Professor of Law, Antonin Scalia Law School, George Mason University; Ph.D., Yale University, Department of Economics; J.D., Stanford University School of Law; M.A., Yale University, Department of Economics.

‡ Professor of Law, Courtesy Professor of Economics, and Director of the Law & Economics Program, Texas A&M University School of Law; Ph.D., Boston College, Department of Economics; J.D., George Mason University School of Law; M.A., Boston College, Department of Economics. We thank Wei-Cheng Chen, Claudine Desrieux, John Shahar Dillbary, Nuno Garoupa, Emma Harrington, Bruce Kobayashi, Suhyeon Oh, Georg Ringe, Ludivine Roussey, Abraham Wickelgren and participants at the European Law and Economics Association Annual Meeting, the Canadian Law and Economic Association Annual Meeting, the American Law and Economics Association Annual Meeting, Southern Economic Association Annual Conference, Paris II CRED-HEC Law, Economics, and Regulation Conference, and the Manne Forum at George Mason University for their valuable feedback. Kendall Alford, Josh Hartt, Nima Kaffi and Abhinav Mishra provided excellent research assistance. We gratefully acknowledge the editors of the *New York University Law Review* for their meticulous editorial work.

INTRODUCTION

Imagine two suspects who are identical in every respect except their race. They are brought before a prosecutor for plea-bargaining. In a perfectly impartial world, where prosecutors and judges are completely blind to race,[1] the two identical suspects face the same plea-bargain process and expect the same outcome[2] since they are assumed to be identical but for their race, which is not observable to the prosecutor and the judge. Building on this reasoning, scholars and policymakers often support reforms such as implicit-bias training,[3] and laws that ban the use of information such as prior convictions, ethnic origin, and so on in hopes that eliminating access to potentially prejudicial data will reduce discrimination.[4]

---

[1] Such an ideal world does not exist. *See, e.g.*, Ram Subramanian, Léon Digard, Melvin Washington II & Stephanie Sorage, *In the Shadows: A Review of the Research on Plea Bargaining*, VERA INST. OF JUST. 24–31 (2020), https://www.vera.org/downloads/publications/in-the-shadows-plea-bargaining.pdf [https://perma.cc/M4NQ-MA4M] (reviewing potential biases in the plea-bargaining process).

[2] *See infra* Section II.B.2 (reviewing differences in plea-bargaining outcomes based on race).

[3] *See*, *e.g.*, Elayne E. Greenberg, *Unshackling Plea Bargaining from Racial Bias*, 111 J. CRIM. L. & CRIMINOLOGY 93, 134 (2020).

[4] *See*, *e.g.*, J.J. Prescott & Sonja B. Starr, *Expungement of Criminal Convictions: An Empirical Study*, 133 HARV. L. REV. 2460, 2463 (2020) ("Perhaps the policy levers with the greatest theoretical potential to improve reentry outcomes are laws that allow criminal conviction

But this line of reasoning misses something crucial: Even when courts and prosecutors are race-blind, discrimination outside the courtroom can still affect bargaining outcomes within. If, for example, police disproportionately arrest minority defendants—as patterns widely documented across jurisdictions suggest[5]—minority individuals in plea negotiations will hold worse "outside options"—what they expect to receive if plea negotiations fail. And because bargaining outcomes depend on outside options—what awaits the defendant if negotiations break down—individuals with worse outside options are more likely to accept worse plea-bargaining outcomes. Therefore, disparate impact may persist despite the complete impartiality of prosecutors and judges. In this sense, discrimination is contagious—it can permeate from one institution to another, even when the latter operates with complete impartiality.

Discrimination undermines fundamental principles of justice, denying individuals opportunities and resources based on their membership in a particular group instead of individual merit. Discrimination has been frequently observed in the hiring process,[6] wages,[7] housing,[8] education,[9] healthcare,[10] and the criminal justice system.[11] Legal philosophers argue that discrimination is wrongful because it demeans individuals[12] and inflicts

---

records to be wholly expunged or, at least, sealed from public view."); Amanda Agan & Sonja Starr, *The Effect of Criminal Records on Access to Employment*, 107 AM. ECON. REV. 560, 560, 563 (2017) (demonstrating how criminal record inquiries on job applications prevents employment for individuals with records).

5 *See, e.g.*, Aline Ara Santo Carvalho, Táhcita Medrado Mizael & Angelo A. S. Sampaio, *Racial Prejudice and Police Stops: A Systematic Review of the Empirical Literature*, 15 BEHAV. ANALYSIS PRAC. 1213, 1214 (2021) (surveying studies finding police disproportionately stop and use unnecessary force against Black and Hispanic individuals).

6 *See, e.g.*, Marianne Bertrand & Sendhil Mullainathan, *Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination*, 94 AM. ECON. REV. 991 (2004); Lincoln Quillian, Devah Pager, Ole Hexel & Arnfinn H. Midtbøen, *Meta-Analysis of Field Experiments Shows No Change in Racial Discrimination in Hiring Over Time*, 114 PROC. NAT'L ACAD. SCI. 10,870, 10,870 (2017) (showing that "levels of discrimination remain largely unchanged, at least at the point of hire").

7 *See, e.g.*, Francine D. Blau & Lawrence M. Kahn, *The Gender Wage Gap: Extent, Trends, and Explanations*, 55 J. ECON. LITERATURE 789 (2017).

8 *See, e.g.*, DOUGLAS MASSEY & NANCY DENTON, AMERICAN APARTHEID: SEGREGATION AND THE MAKING OF THE UNDERCLASS (1993).

9 *See, e.g.*, JEANNIE OAKES, KEEPING TRACK: HOW SCHOOLS STRUCTURE INEQUALITY (2005).

10 *See, e.g.*, Leiyu Shi, Chien-Chou Chen, Xiaoyu Nie, Jinsheng Zhu & Ruwei Hu, *Racial and Socioeconomic Disparities in Access to Primary Care Among People with Chronic Conditions*, 27 J. AM. BD. FAM. MED. 189 (2014) (finding racial disparities in access to medical care).

11 *See, e.g.*, Nazgol Ghandnoosh, *One in Five: Ending Racial Inequity in Incarceration*, SENT'G PROJECT (2023), https://www.sentencingproject.org/app/uploads/2024/02/One-in-Five-Ending-Racial-Inequity-in-Incarceration.pdf [https://perma.cc/S7AL-ATCS]; Carvalho, Mizael & Sampaio, *supra* note 5.

12 *See* DEBORAH HELLMAN, WHEN IS DISCRIMINATION WRONG? 29–33 (2008).

social stigmas on individuals and their groups.[13] Researchers have identified various social harms of discrimination. First and foremost, discrimination insults, hurts, and stigmatizes disfavored groups, making them feel inferior and angry and potentially leading to violence.[14] Discrimination has been found to severely affect the mental health of its victims.[15] Discrimination is also harmful to society in multiple ways. First, as an economic matter, discrimination in the labor market has been found to result in a reduction in aspirations, energy, and productivity.[16] Second, differences in wages for jobs of comparable worth distort economic decisions and result in social inefficiency. More generally, when discrimination is widespread, it can lead discriminated individuals to accept limiting roles that constrain their flourishing, resulting in lost productivity.[17] Research has shown that discrimination can also lead to disharmony and social violence,[18] as well as increases in crime rates.[19]

Recognizing and understanding contagious discrimination is thus crucial. Policymakers should recognize that downstream courtroom reforms alone—for example, debiasing trainings for prosecutors or judges—are insufficient. Because discriminatory practices by police or other upstream actors can cascade through the system, eliminating downstream disparities requires attention to upstream bias. Addressing bias at its source is therefore crucial, and in some cases, tackling upstream discrimination may also be more cost-effective than piecemeal reforms further along the chain.

This Essay also advances the broader social-scientific study of discrimination.[20] Economists typically emphasize two mechanisms: "taste-based discrimination,"[21] in which decision makers harbor preferences against a group, and "statistical discrimination,"[22] in which they rely on group characteristics as proxies for unobservable traits such as productivity. Both approaches assume that discrimination originates with the actor who directly observes the protected characteristic. Our theory demonstrates something

---

13 *See* IYIOLA SOLANKE, DISCRIMINATION AS STIGMA: A THEORY OF ANTI-DISCRIMINATION LAW 17–38 (2017).

14 Larry Alexander, *What Makes Wrongful Discrimination Wrong? Biases, Preferences, Stereotypes, and Proxies*, 141 U. PA. L. REV. 149, 162 (1992).

15 *See, e.g.*, Yvonne Lei et al., *Discrimination and Subsequent Mental Health, Substance Use, and Well-Being in Young Adults*, 148 PEDIATRICS e2021051378, 1, 5 (2021).

16 Alexander, *supra* note 14, at 162.

17 *Id.* at 165.

18 *Id.* at 198.

19 *See* Murat C. Mungan, *Statistical (and Racial) Discrimination, "Ban the Box," and Crime Rates*, 20 AM. L. ECON. REV. 512, 513–14 (2018) (showing that statistical discrimination in the labor market can lead to an increase in crime rate in the disadvantaged groups, and, under some additional conditions, an increase in the overall crime rate).

20 *Infra* Part I.

21 *Infra* Section I.A.1.

22 *Infra* Section I.A.2.

different: discrimination can persist even when these downstream actors do not observe race, gender, or religion at all.

Our theory also departs from recent work by prominent sociologists who urged economists to move beyond individual-level analysis toward institutional or organizational discrimination.[23] While our theory responds to their call, it remains distinct from organizational or institutional discrimination that sociologists have already carefully studied. Rather than focusing on institutions and laws that perpetuate existing disparities,[24] our theory highlights a new mechanism by which discrimination can spread across different institutions.

Part I reviews existing theories of discrimination in economics and sociology, including taste-based, statistical, and organizational discrimination. It concludes by highlighting the novelty of our theory in contrast to existing theories. Part II introduces our new theory of contagious discrimination through a simple illustration using job search as a motivating example, demonstrating that the phenomenon of contagious discrimination can occur in various institutional settings. Part III develops a more complex model of contagious discrimination in the criminal justice system, focusing particularly on the plea-bargaining process and showing how discriminatory conduct by the police can influence the plea-bargaining process, even when prosecutors and judges are race-blind. Part IV discusses the implications uncovered by our new theory of discrimination for policymakers in criminal justice, emphasizing the importance of (1) ensuring that defendants do not hold incorrect perceptions of police discrimination and (2) avoiding evaluations of police officers based on the final case dispositions of their arrest decisions. We discuss potential reforms in plea-bargaining to facilitate cooperative bargaining between prosecutors and defense lawyers and to reduce opacity in the system by creating a transparent sentence reduction standard to limit prosecutorial discretion.

I
A REVIEW OF THEORIES OF DISCRIMINATION

This Part reviews theories of discrimination in economics (Section I.A) and sociology (Section I.B). Section I.C discusses these existing theories and provides the motivation for our new theory of contagious discrimination.

---

23 *See* Mario L. Small & Devah Pager, *Sociological Perspectives on Racial Discrimination*, 34 J. ECON. PERSPS. 49, 52 (2020). Legal scholars have also provided important contributions to our understanding of discrimination. *See, e.g.*, TARUNABH KHAITAN, A THEORY OF DISCRIMINATION LAW, 107–58 (2015) (reviewing non-economic theories of discrimination).

24 Small & Pager, *supra* note 23, at 53–60.

A. Economic Theories of Discrimination

Economic theories of discrimination at their inception focused on employment discrimination,[25] but these models can be generalized to other areas, such as law enforcement and other aspects of the criminal justice system which will be reviewed in this Section.[26]

Discrimination refers to decision-making that treats individuals worse (in terms of opportunities, outcomes, or burdens) because of their membership in a socially salient, often observable group—for example, race or gender—rather than because of relevant individual conduct or qualifications. Economists identify two main types of discrimination: taste-based discrimination and statistical discrimination. These differ in their underlying causes. Taste-based discrimination relies on intrinsic preferences, while statistical discrimination arises from optimization under incomplete information. To illustrate, consider a scenario where female job candidates receive more interview opportunities than male candidates at a firm. This discrepancy would indicate taste-based discrimination if the employer prefers hiring females and therefore invites more female candidates for interviews. In contrast, it could be a case of statistical discrimination if the employer, lacking sufficient information about the candidates, offers more opportunities to females not due to personal preference for females, but because the employer believes females are, on average, more productive than males for the job in question. In this way, taste-based discrimination is based on characteristics that are economically irrelevant to the decision, whereas statistical discrimination relies on self-interested optimization, using group characteristics as imperfect signals of relevant traits that are not directly observable. Importantly, taste-based discrimination and statistical discrimination are not mutually exclusive; discriminatory practices can arise from both simultaneously.[27]

1. Taste-Based Discrimination

Taste-based racial discrimination was first theorized by Gary Becker.[28]

---

25 *See, e.g.*, GARY S. BECKER, THE ECONOMICS OF DISCRIMINATION (1957); Kenneth J. Arrow, *The Theory of Discrimination*, *in* DISCRIMINATION IN LABOR MARKETS 3 (Orley Ashenfelter & Albert Rees eds., 1973); Edmund S. Phelps, *The Statistical Theory of Racism and Sexism*, 62 AM. ECON. R. 659, 659–61 (1972) (noting that Gary Becker "originated" the economic theory of discrimination).

26 *See, e.g.*, John Knowles, Nicola Persico & Petra Todd, *Racial Bias in Motor Vehicle Searches: Theory and Evidence*, 109 J. POL. ECON. 203, 205 (2001) (applying Becker's model for testing for discrimination to motor vehicle searches).

27 *See infra* note 133 and accompanying text (citing to studies that distinguish the two theories).

28 *See* BECKER, *supra* note 25 (conceptualizing how employers can have a "taste for discrimination").

Becker's theory posits that individuals may prefer to associate with a particular group based on their tastes or preferences. In other words, they may have a "taste" for discrimination against certain groups and in favor of others. Quite importantly, in this framework, taste is not related to other relevant factors such as productivity in the context of employment or the probability of committing a crime in the context of law enforcement.[29] There is a large empirical literature that attempts to identify taste-based discrimination across different dimensions such as race,[30] ethnic origin,[31] and gender.[32] While taste-based discrimination is often entangled with statistical discrimination, a recent study attempts to isolate taste-based discrimination with respect to race, gender, and age by making participants' payoffs independent of their choices in the game-show *Big Brother*.[33] For discrimination against women, the cross-culture study finds (1) strongly significant discrimination in German, Italian and British versions of *Big Brother*, (2) weakly significant discrimination in the Spanish version, and (3) no significant discrimination in the Australian and Brazilian versions.[34] For racial discrimination against non-white people, the study suggests that strong significant effects are observed in Germany and Italy and a weak significant effect is observed in the UK.[35] Finally, for age discrimination, the study finds significant preferences for younger contestants in Italy only.[36] These studies demonstrate that taste-based discrimination is a real phenomenon encountered frequently and that its severity depends on cultural and other factors.

A second part of Becker's original contribution is the prediction that an

---

29 Mungan, *supra* note 19, at 516.

30 *See, e.g.*, Devah Pager, *Are Firms That Discriminate More Likely to Go Out of Business?*, 3 SOCIO. SCI. 849, 852 (2016) (finding that twenty-four percent of employers showed evidence of discrimination against Black people on the labor market).

31 Magnus Carlsson & Dan-Olof Rooth, *Revealing Taste-Based Discrimination in Hiring: A Correspondence Testing Experiment with Geographic Variation*, 19 APPLIED ECON. LETTERS 1861, 1864 (2012) (finding that "in municipalities with more negative attitudes, especially for low-skilled occupations, applicants with a typical Middle Eastern name are discriminated to a larger extent").

32 *See, e.g.*, Hadar Gafni et al., *Gender Dynamics in Crowdfunding (Kickstarter): Evidence on Entrepreneurs, Backers, and Taste-Based Discrimination*, 25 REV. FIN. 235, 265 (2021) ("[A]s they gain experience on the platform, female backers become agnostic to gender, while male backers maintain their discriminatory behavior. . . . [M]ale backers' tendency to back male entrepreneurs was partly due to [taste-based discrimination].").

33 *See* Tom Lane, *Get Her off My Screen: Taste-Based Discrimination in a High-Stakes Popularity Contest*, 71 OXFORD ECON. PAPERS 548, 558–59 (2019) ("An important motivation for the study comes from the fact that, as the audience do not have own material payoffs contingent on their votes, the setting is one where statistical discrimination cannot occur and any discrimination identified is therefore necessarily taste-based.").

34 *Id.* at 559.

35 *Id.*

36 *Id.*

employer practicing taste-based discrimination would gradually lose its competitiveness: Because the employer's preference includes factors that are not related to productivity, the employer is not able to optimize as well as its competitors that do not have a taste for discrimination.[37] This powerful market-based argument is empirically examined in a recent paper on racism in the labor market.[38] This study found that discriminatory businesses are more than twice as likely as non-discriminatory businesses to go out of business.[39] Although the study finds "a strong and significant association between discrimination and firm survival,"[40] the author cautions that it is premature to conclude that discrimination itself *caused* an employer to go out of business in accordance with Becker's prediction:

> The present results are consistent with Becker's theory of discrimination, although they do not allow us to draw specific conclusions about the mechanisms by which discriminatory decision-making is associated with firm survival. It is possible, for example, that the kinds of employers who discriminate against racial minorities are also those who make poor choices in other areas of business management. As one example, in interviews with employers about hiring practices, one of the most common expressions used to describe their hiring philosophy is a reliance on a "gut feeling" (internal citations omitted). Employers who use intuition—rather than evidence, or more systematic forms of decision-making—may also be led astray in their capital investments, their financial management, and numerous other domains (internal citation omitted). We cannot, therefore, conclude that it is directly because of the costs of discrimination that employers in our sample appear more likely to have been driven out of business.[41]

Becker's hypothesis that, in the long run, market forces will correct taste-based discrimination presumes that the discriminatory act is under the influence of market forces. But not all types of taste-based discrimination are subject to correction by the market, which Becker also acknowledges.[42] Consider, for example, government actors such as the police. The free market will not correct a police officer's discriminatory conduct because the police officer is funded by tax money—police departments do not go out of

---

37 BECKER, *supra* note 25, at 35–37.

38 *See* Pager, *supra* note 30.

39 *Id.* at 852 (finding that "17 percent of nondiscriminatory establishments had failed by 2010, relative to 36 percent of those that did discriminate" with $p < -0.05$).

40 *Id.*

41 *Id.* at 855.

42 *See, e.g.*, Kevin M. Murphy, *How Gary Becker Saw the Scourge of Discrimination*, CHI. BOOTH REV. (June 15, 2015), https://www.chicagobooth.edu/review/how-gary-becker-saw-the-scourge-of-discrimination [https://perma.cc/2VWX-4U5X] ("Becker argued that in terms of schooling, the United States had routinely discriminated, because the competitive forces that govern the market are not present in state-provided elementary and secondary education.").

business if they are not as competitive in making arrests.[43] In such cases, government regulations may be necessary to correct such discriminatory behavior.

2. Statistical Discrimination

Statistical discrimination is based on optimization under incomplete information. With statistical discrimination, a person's observable characteristics, such as their gender and race, are used as predictors of their unobservable qualities such as productivity. An employer, for example, may have no intrinsic or market-driven preference for any race or gender but nevertheless give preferential hiring to a certain group because the employer believes that group to be more productive on average and cannot observe productivity from the job-screening process.

A substantial amount of research on statistical discrimination[44] can be traced to the seminal articles by Professors Phelps[45] and Arrow.[46] Professor Phelps presumes differences in productivity across groups, whereas Professor Arrow examines the mechanism that entrenches such differences.[47] To illustrate, in the context of employment, Phelps would simply assume that people in different groups have different levels of productivity. By contrast, Professor Arrow explains that such differences may arise from self-fulfilling stereotypes: A belief that the discriminated minority group is less productive than the majority group can cause the minority group to choose to invest less in developing their skills, thus entrenching the existing belief that they are less productive.[48] A more recent study demonstrates that statistical discrimination can take place even in the absence of any divergent belief in cross-group productivity if employers are better able to screen one group than another.[49]

---

43 *See* Jeremy Travis, *Measuring What Matters; Part Two: Developing Measures of What the Police Do*, NAT'L INST. JUST. (Nov. 1997), https://www.ojp.gov/pdffiles/167255.pdf#:~:text=When%20we%20worked%20for%20police,that%20the%20bottom%20line%20in [https://perma.cc/Z4H8-92Z6].

44 *See* Hanming Fang & Andrea Moro, *Theories of Statistical Discrimination and Affirmative Action: A Survey*, *in* 1 HANDBOOK OF SOCIAL ECONOMICS 133, 133–200 (Jess Benhabib, Alberto Bisin & Matthew O. Jackson eds., 2011) (reviewing the literature on statistical discrimination).

45 Phelps, *supra* note 25.

46 ARROW, *supra* note 25.

47 In economic parlance, the two sub-categories of statistical discrimination differ in that Phelps assumes *exogenous* differences between groups whereas Arrow *endogenizes* the differences. Exogenous differences are determined outside the model, which takes indifferences as given. Endogenous differences, by contrast, are determined within the model.

48 *See, e.g.*, Stephan Coate & Glenn C. Loury, *Will Affirmative-Action Policies Eliminate Negative Stereotypes?*, 83 AM. ECON. REV. 1220, 1221–22 (1993) (suggesting that statistical discrimination could discourage minorities from investing in skills by leading them to believe that such investments would not be fully rewarded).

49 Bradford Cornell & Ivo Welch, *Culture, Information, and Screening Discrimination*, 104 J.

Professor Arrow's model of endogenous statistical discrimination has been adapted to the criminal law context, showing that even when groups have identical characteristics, statistical discrimination can result from "self-fulfilling expectations."[50] For example, even if there are no intrinsic differences in criminality between Black people and white people, a belief that Black people are more likely to commit crimes than white people would result in the stigmatization of Black people. And with stigmatization, Black people have fewer employment opportunities than white people and are thus more likely to choose to commit crimes due to unemployment.[51] This feedback loop then entrenches the initial belief that Black people are more likely to commit crimes.

Because statistical discrimination involves optimization under incomplete information, it may be interpreted as being less harmful than having a taste for discrimination and exclusion.[52] This view has been contested on the utilitarian grounds that statistical discrimination may be even more harmful and harder to eliminate through legal design than taste-based discrimination.[53] This difficulty is because eliminating statistical discrimination would require undoing self-perpetuating beliefs, which is typically more challenging than using other incentives to counteract taste-based preferences to encourage inclusive behavior.[54]

Moreover, beliefs which form the basis of statistical discrimination may simply be inaccurate. For example, an employer's belief that members of a certain race are generally better at communication than others may be wrong or outdated. An important strand of literature in labor economics focuses on statistical discrimination in the labor market based on such beliefs.[55] It argues employers' positive experiences with discriminated minority groups will result in the correction of such erroneous beliefs over time.[56]

POL. ECON. 542, 545 (1996).

50 Eric Rasmusen, *Stigma and Self-Fulfilling Expectations of Criminality*, 39 J.L. & ECON. 519, 519 (1996).

51 *See, e.g.*, *id.*; Patricia Funk, *On the Effective Use of Stigma as a Crime-Deterrent*, 48 EUR. ECON. REV. 715, 716 (2004) (explaining the effects of stigma on employment opportunities); Murat C. Mungan, *A Generalized Model for Reputational Sanctions and the Relevance of the Interactions Between Legal and Reputational Sanctions*, 46 INT'L REV. L. & ECON. 86, 87 (2016) (same).

52 *See* Murat C. Mungan, *Discrimination and Deterrence with Enforcer Liability*, 22 AM. L. ECON. REV. 303, 304–07 (2020).

53 *See id.*

54 *See id.*

55 *See generally* Joseph G. Altonji & Charles R. Pierret, *Employer Learning and Statistical Discrimination*, 116 Q.J. ECON. 313 (2001) (providing empirical evidence supporting the theory of statistical discrimination in the labor market); Henry S. Farber & Robert Gibbons, *Learning and Wage Dynamics*, 111 Q.J. ECON. 1007 (1996) (same); Amy Farmer & Dek Terrell, *Discrimination, Bayesian Updating of Employer Beliefs, and Human Capital Accumulation*, 34 ECON. INQUIRY 204 (1996) (same).

56 *See, e.g.*, Farmer & Terrell, *supra* note 55.

This optimistic view that erroneous beliefs will over time disappear has since been challenged.[57] A recent survey study of employers shows that, while employers update their expectations for *individual* workers, they do not update their *group* expectations.[58]

B. Sociological Theory of Discrimination

In a recent article, two prominent sociologists, Professors Mario Small and Devah Pager, argue that taste-based and statistical discrimination, with their emphasis on individual decision-making, are too narrow in scope:

> [A] substantial body of evidence suggests that limiting the study of discrimination to the actions of potentially prejudiced individuals dramatically understates the extent to which people experience discrimination; understates the extent to which discrimination may account for social inequality; and understates the extent to which discrimination may play a role in markets for labor, credit, and housing, as well as in other contexts.[59]

Professors Small and Pager highlight important studies in sociology that demonstrate that discrimination can occur without regard to individuals' intentions.[60]

Institutional discrimination occurs when organizations or laws perpetuate a pattern of differential treatment, not through the individual institutional players' taste-based or statistical discrimination, but as a consequence of preexisting disparities.[61] Professors Small and Pager provide an example of an organization's decision to lay off employees based on their tenure. This race-neutral mechanism could result in a disparate impact on racial minorities and women, as many organizations have only recently begun to include large numbers of these minority groups.[62] Applying the race-neutral mechanism to layoffs, in light of the preexisting disparity (where racial minorities and women are more likely to have shorter tenures) would therefore lead to a higher number of layoffs among minorities.[63]

Similarly, race-neutral laws can perpetuate preexisting disparities, even if the institutional actors applying these laws do not discriminate based on taste or statistical optimization. Professors Small and Pager cite laws that

---

57 *See, e.g.*, Devah Pager & Diana Karafin, *Bayesian Bigot? Statistical Discrimination, Stereotypes, and Employer Decision Making*, 621 ANNALS AM. ACAD. POL. & SOC. SCI. 70, 89–90 (2009) ("Where employers may update their expectations regarding individual workers, these experiences do not seem to have noticeable effects on their attitudes about the group as a whole.").

58 *Id.*

59 Small & Pager, *supra* note 23, at 52.

60 *Id.*

61 *Id.* at 53–54.

62 *Id.* at 54.

63 *Id.*

disenfranchise imprisoned felons or people who have been convicted of a felony as an example. Given that Black Americans are disproportionately convicted of felonies, this race-neutral policy results in the disenfranchisement of a disproportionate number of Black Americans.[64] They further explain that many of these seemingly race-neutral disenfranchisement laws were enacted after Black Americans were granted the right to vote.[65]

C. Discussion and Motivation for Our Theory

Taste-based discrimination and statistical discrimination are often intertwined. Consider the well-known experimental finding that employers are fifty percent more likely to give callbacks to candidates with white-sounding names compared to those with Black-sounding names, despite equivalent qualifications on their resumes.[66] However, this phenomenon may take place because some employers have a taste for white candidates, believe that white employees are likely to be more productive given the same resume than non-white employees, or a combination of *both*.

We concur with Professors Small and Pager that these two economic theories of discrimination cannot fully explain all situations where discrimination occurs. We also agree that seemingly neutral institutions (organizations and laws) can perpetuate existing disparities.

However, there is a gap in the literature regarding how individual-based economic theories of discrimination relate to discrimination at the institutional level. This motivates us to propose a theory to fill this gap: discrimination arising from individual interactions that occur during negotiations, bargaining, and other instances of surplus splitting.[67]

II
CONTAGIOUS DISCRIMINATION THROUGH BARGAINING: AN ILLUSTRATION AND REVIEW OF THE BARGAINING LITERATURE

Our theory shows that disparate outcomes can arise even when the decision maker is color-blind. Under our theory, the way discrimination spreads is neither through taste-based nor statistical dynamics, as the decision maker does not need to observe a person's characteristics such as their race. Our theory also differs from institutional discrimination in that it is not laws or organizations that perpetuate discrimination at the institutional

64 *Id.* at 59.

65 *Id.* ("[T]hese laws rose dramatically in number and scope after the Civil War, following Reconstruction and the ratification of the Fifteenth Amendment, which gave African-Americans the right to vote[.]" (citation omitted)).

66 Bertrand & Mullainathan, *supra* note 6, at 992.

67 Robert Cooter & Thomas Ulen, LAW AND ECONOMICS 74–86 (6th ed. 2011).

level. Instead, we show that the interactive bargaining process can inherently perpetuate disparity. When negotiations involve individuals who are disproportionately affected, even if the other side of the negotiation does not discriminate, those individuals may still end up with worse outcomes due to limited outside options. In other words, discriminatory practices outside the negotiation can infiltrate otherwise neutral institutions.

This Part begins with a simple illustration of our theory of contagious discrimination in the context of job search. Because the theory of contagious discrimination relies on key properties of outcomes obtained through the bargaining processes, Section II.B reviews the bargaining literature in general with an emphasis on these key properties.

A. Illustration of Contagious Discrimination

Imagine a person is searching for a job. She receives an offer for a job that would pay her $5,000 per month for five years. Should she stop her job search and accept the job offer? The answer to this question depends, among other things, on how soon she anticipates she will receive another job offer and how likely it is that the next job offer will provide her with a more attractive package.

The prospect of receiving another job offer and the likelihood of that job offer being superior both affect her outside option, which refers to what she expects to get if she turns the current offer down. The better her outside option, the more likely she is to decline the current offer; the employer must make a stronger offer to her to induce her to accept. In other words, a superior outside option strengthens her bargaining position relative to the employer sitting on the other side of the bargaining table. By contrast, the same analysis suggests that, all else equal, discrimination in the labor market[68] weakens the outside option of a discriminated job seeker relative to an otherwise identical employee who does not face discrimination.

Thus, a worse outside offer implies that the discriminated employee is more likely to accept a worse job offer now—if you do not believe you will receive another job offer soon, the relative value of the current offer is greater.[69] Discriminatory practices by future employers thus weaken the bargaining position of the individual, inducing her to agree to a worse bargaining outcome in the present. This is true even if the current employer with whom she negotiates does not discriminate at all.

---

68 One example of this type of discrimination is reduced probability of employment contract. *See, e.g.*, Patrick Kline, Evan K. Rose & Christopher R. Walters, *Systemic Discrimination Among Large U.S. Employers*, 137 Q.J. ECON. 1963, 1965 (2022) (showing that candidates with distinctively Black names are 2.1% less likely to receive employer contract in comparison to similarly positioned candidates with distinctively white names).

69 An underlying assumption is that accepting the job offer is better than not doing any job.

B. A Brief Review of the Bargaining Literature

The foregoing thought experiment in the job search context illustrates the intuition behind our theory: Discrimination anticipated in the future weakens an individual's outside option or bargaining position in the present. The weaker outside option of an individual anticipating discrimination implies a comparatively less favorable bargaining outcome for her. As a result, such an individual is willing to accept a lesser offer than an otherwise identical individual who is not discriminated against.

This Section reviews theoretical and empirical bargaining studies in the academic literature which support this prediction.

1. Bargaining Theories and Outside Options

Bargaining theories model how parties negotiate the division of surplus from trade.[70] This surplus is measured by the difference between the value of a successful bargaining agreement and the sum of the parties' outside option values that would be obtained upon negotiation failure.[71] Therefore, outside options directly affect the size of the surplus.

In standard bargaining problems, this surplus is split between two parties. Which party receives what proportion of the surplus can depend on many factors, including the bargaining power of the two parties.[72] To illustrate, consider a numerical example: A seller values a car at \$500[73] and a buyer values the car at \$1500. In this case, a successful bargaining outcome creates a surplus of \$1000 in comparison to a failed negotiation.[74] This is because the buyer's surplus is the difference between what she pays for the car (denoted p) and her valuation of the car (i.e., \$1500), which equals $\$1500 - \$p$; and the seller's surplus equals the amount she gets from the sale of the car and his valuation of the car, i.e., $\$p - \$500$. The total

70 *See, e.g.*, Cooter & Ulen, *supra* note 67, at 74–80.

71 *Id.*

72 There are many models studying how this split is or can be determined. One quintessential example is the bargaining theory proposed by the Nobel Prize laureate Professor John Nash. *See generally* John F. Nash, Jr., *The Bargaining Problem*, 18 ECONOMETRICA 155 (1950) (modeling Nash's bargaining theory). Another is Professor Rubinstein's theory, which we describe below. *See generally* Ariel Rubinstein, *Perfect Equilibrium in a Bargaining Model*, 50 ECONOMETRICA 97 (1982) (modeling Rubinstein's bargaining theory). In many applications, it is convenient to take the characteristics of the two parties as given, and describe how they determine the share of the surplus received by the two parties through a single parameter. *See generally* Murat C. Mungan, *Reverse Payments, Perverse Incentives*, 27 HARV. J.L. & TECH. 1 (2013) (using the Nash bargaining solution in the context of settlement negotiations to determine the division of surplus between two parties based on their relative bargaining power).

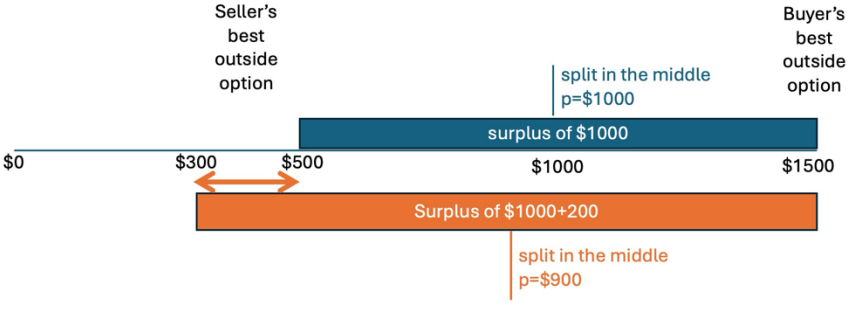
73 This could, for example, be the next best offer on the market for the car.

74 A successful trade leads to the car going from the seller to the buyer, who values the car \$1000 more than the buyer. The exchange thus creates a value of \$1000. The price paid by the buyer to the seller is a transfer: from society's perspective, there is no net gain or loss associated with a transfer payment.

surplus for the two parties is thus $(\$1500 - \$p) + (\$p - \$500) = \$1000$.

The amount of the surplus obtained by each party can be described by a parameter, which we denote α following the convention in the literature.[75] This parameter measures the share of the surplus that one of the parties, say the seller, gets from the sale of the car.[76] For illustration purposes, consider the case where the two sides split their surplus evenly, in which case $\alpha = 0.5$. Then, the buyer and seller will both obtain $\$1000 * 0.5 = \$500$ of the surplus. This is achieved through a sales price of $1000, which splits the surplus down the middle.[77] This is demonstrated in the graph below in blue (top half) where the surplus of $1000 is evenly split between the seller and buyer—$500 of the surplus goes to the seller and $500 goes to the buyer at a price of $1000. The buyer pays $500 less than her best outside option and the seller sells for $500 more than her best outside option.

FIGURE 1. ILLUSTRATION OF BARGAINING THEORY



A result that emerges in this setting is that a worsening in the seller's outside option leads to a reduction in the sale price.[78] To illustrate, suppose that the seller's valuation of the car is reduced from $500 to $300.[79] This

75 *See* Ken Binmore, Ariel Rubinstein & Asher Wolinsky, *The Nash Bargaining Solution in Economic Modelling*, 17 RAND J. ECON. 176, 186 (1986).

76 Professor Nash did not introduce α in his 1950 publication, but it has become an often-used parameter in the applied economics literature.

77 As an illustration of unequal bargaining power, suppose the seller has stronger bargaining power. In this case, the seller captures a larger portion of the surplus—for example, $700, which corresponds to a bargaining weight of $\alpha = 0.7$.

78 This general result can be shown mathematically. All else being equal, a worsening in the seller's outside option by $\$x$ expands the surplus by $\$x$. However, the seller only gets to enjoy her part, $\alpha * \$x$, of the surplus and the buyer gets to enjoy the remaining $(1 - \alpha) * \$x$. Price therefore decreases by $(1 - \alpha) * \$x$ since the buyer's outside option is unchanged.

79 This happens, for example, when the other best offeror of $500 withdraws the offer, leaving the seller now with the second-best offer of $300 (numbers are made up for the purpose of illustration).

increases the surplus from $1000 to $1200.[80] An equal split of the surplus is now achieved through a sale in which the seller gets $600 above his valuation, at a price of $300 + $600 = $900. Thus, as suggested, a worsening in the seller's outside option leads to a lower and worse sales price for the seller.[81] This is illustrated in orange (bottom half) of Figure 1 above. A worsening of the seller's best outside option from $500 to $300 expands the total surplus from $1000 to $1200, with an increase of $200. However, only part (half when they have equal bargaining power) of that surplus accrues to the seller with the other half going to the buyer. The final price is therefore lower at $900—it is worse for the seller who would prefer to sell at the previous price of $1000.

The intuition behind this result is simple. A worse outside option increases the total surplus for the two parties. However, only a part of that surplus (in our specific example, half of it) goes to the party who now has a worse outside option. Therefore, the overall effect is a reduction in the value obtained by the party who has a worse outside option and an increase of the same size in the benefit of the other party.

The bargaining model we have outlined here assumes that the parties' bargaining power, reflected by α, is taken as given. In a highly influential article, Professor Ariel Rubinstein proposed a way to explain how the two parties' bargaining powers may be determined.[82] In this article, the patience of the parties—captured through how much they discount future benefits and costs relative to benefits and costs today—determines their relative bargaining powers.[83] For our purposes, the key intuition and predictions that emerge in the bargaining model we have outlined and the Rubinstein Bargaining settings are the same: The worsening of a party's outside option expands the surplus, but only part of that surplus goes to the party with a worse outside option.

It is important to note that scholars have criticized the application of these rational models. For example, then-Professor Stephanos Bibas, critiquing bargaining theory in the setting of plea-bargains, points out that additional factors such as lawyering and psychological biases can also distort bargaining outcomes.[84] Moreover, the assumption of complete information (all information is observable to players) has been relaxed, leading to an

---

80 The surplus is now $1500 − $300 = $1200.

81 The result also holds if the buyer has a worse outside option. Suppose the buyer values the car more at $2100 and cannot get any substitutes on the market for a lower price. This corresponds to a surplus of $1600. Equal split of the surplus would result in a sale price of $500 + $800 = $1300—a worse price for the buyer, who now has to pay more.

82 *See* Rubinstein, *supra* note 72 at 97–98.

83 *See id.* at 99.

84 *See* Stephanos Bibas, *Plea Bargaining Outside the Shadow of Trial*, 117 HARV. L. REV. 2464, 2466–527 (2004) (reviewing and critiquing shadow-of-trial studies).

important strand in the literature.[85] Some of these highlight the importance of information on not only bargaining outcomes but also the likelihood of reaching an agreement,[86] while others examine other important factors that may influence bargaining behavior such as fairness considerations[87] and strategic behavior.[88] However, the relationship between parties' outside options and the eventual terms of bargaining that they agree to are not necessarily affected by these considerations, unless they interact in unexpected ways. Whether the relationship holds true in real life is an empirical question, and thus we briefly review the empirical literature on bargaining next.

2. Empirical Studies

A separate strand of the literature focuses on empirical studies of bargaining in different fields.[89] In these studies, scholars analyze available data from real world settings to study, among other things, the effect of outside options on bargaining outcomes. It is important to note the difficulty of studying outside options in real life settings, where options are difficult to quantify, and it is harder to control for other factors that may affect bargaining outcomes. Nevertheless, there have been empirical studies that are designed to circumvent these issues.

First, in a recent study of employment bargaining, Professors Sydnee Caldwell and Oren Danieli used employment data from Germany to develop an outside options index.[90] They found that limited outside options (such as lower geographic mobility), as opposed to discrimination and preferences, accounted for twenty percent of the gender wage gap.[91]

More directly relevant to our theory, studies on plea-bargaining provide evidence that worse outside options held by defendants lead to worse

---

85 Economists refer to this strand as bilateral trade under imperfect information. *See, e.g.*, Ken Binmore, Avner Shaked & John Sutton, *An Outside Option Experiment*, 104 Q.J. ECON. 753, 766 (1989); Vincent P. Crawford & Joel Sobel, *Strategic Information Transmission*, 50 ECONOMETRICA 1431, 1433 (1982); Roger B. Myerson & Mark A. Satterthwaite, *Efficient Mechanisms for Bilateral Trading*, 29 J. ECON. THEORY 265, 267 (1983); Alvin E. Roth & J. Keith Murnighan, *The Role of Information in Bargaining: An Experimental Study*, 50 ECONOMETRICA 1123, 1130 (1982); Richard H. Thaler, *Anomalies: The Ultimatum Game*, 2 J. ECON. PERSPS. 195, 202–05 (1988); Lucian A. Bebchuk, *Litigation and Settlement Under Imperfect Information*, 15 RAND J. ECON. 404, 406 (1984).

86 *See, e.g.*, Bebchuk, *supra* note 85, at 408–09.

87 *See, e.g.*, Thaler, *supra* note 85, at 205; Binmore, Shaked & Sutton, *supra* note 85.

88 *See, e.g.*, Crawford & Sobel, *supra* note 85.

89 *See* Subramanian, Digard, Washington II & Sorage, *supra* note 1 (reviewing plea-bargain studies).

90 Sydnee Caldwell & Oren Danieli, *Outside Options in the Labor Market*, 91 REV. ECON. STUD. 3286, 3287 (2024).

91 *Id.* at 3288.

outcomes.[92] Some have demonstrated that plea-bargaining outcomes are positively correlated with the strength of evidence.[93] Others focus on pretrial detention, demonstrating that defendants held in pretrial detention pled guilty and received harsher sentences than similarly positioned defendants who were not held in pretrial detention.[94] Scholars studying this setting suggest that this is in part because pretrial release provided defendants with better outside options (lower conviction rates) due to less bias, stronger incentives to seek release, and more preparation time for their defense.[95]

Finally, in terms of discrimination and plea-bargaining, Professors Kutateladze, Andiloro, Johnson, and Spohn found that Black and Hispanic defendants are more likely than white defendants to receive a plea offer from prosecutors that includes incarceration, controlling for crime severity, and criminal history.[96] The researchers point to multiple factors that may explain their findings, such as socioeconomic status and the ability to hire private attorneys who help reduce pretrial detention for their clients.[97]

Another study found that Asian American defendants are likely to have more favorable outside options in the form of lower expected trial sentences,[98] and thus Asian Americans may view risk of sentencing more favorably in comparison to Black and Hispanic defendants.

III
CONTAGIOUS DISCRIMINATION IN PLEA BARGAINING

The theory of contagious discrimination focuses on how outside discrimination by certain individuals can trickle down to a discrimination-free bargaining process by worsening the outside option of the individual discriminated against. When this individual later interacts with a non-

92 *See, e.g.*, Subramanian, Digard, Washington II & Sorage, *supra* note 1 (reviewing plea-bargain studies).

93 *See id.* at 19–24 (reviewing recent studies).

94 *See, e.g.*, Will Dobbie, Jacob Goldin & Crystal S. Yang, *The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges*, 108 AM. ECON. REV. 202, 203 (2018); Ellen A. Donnelly & John M. MacDonald, *The Downstream Effects of Bail and Pretrial Detention on Racial Disparities in Incarceration*, 108 J. CRIM. L. & CRIMINOLOGY 775, 805 (2019); *see* Subramanian, Digard, Washington II & Sorage, *supra* note 1, at 11–15 (reviewing the pretrial detention literature).

95 Dobbie, Goldin & Yang, *supra* note 94, at 234 ("[P]retrial release may strengthen a defendant's bargaining position during plea negotiations. . . . While there is no conclusive evidence on this issue, two pieces of evidence suggest that our results are likely driven by changes in a defendant's bargaining position.").

96 Besiki L. Kutateladze, Nancy R. Andiloro, Brian D. Johnson & Cassia C. Spohn, *Cumulative Disadvantage: Examining Racial and Ethnic Disparity in Prosecution and Sentencing*, 52 CRIMINOLOGY 514, 531 (2014).

97 *Id.* at 542.

98 Brian D. Johnson & Sara Betsinger, *Punishing the "Model Minority": Asian-American Criminal Sentencing Outcomes in Federal District Courts*, 47 CRIMINOLOGY 1045, 1064 (2009).

discriminatory actor, her worse outside option weakens her bargaining position, inducing her to accept a worse negotiated outcome than a similarly situated individual who is not externally discriminated against.

The theory of contagious discrimination can be applied to a broader context such as employment, credit applications, rental applications—indeed any interactions involving bargaining or negotiation. This Part, however, applies our theory to the plea-bargain setting to demonstrate important implications that our new theory can help uncover.

Consider, for example, existing scholarship that focuses on institutional design to combat discrimination by reducing biases in the judiciary. For instance, recognizing that society cannot entirely prevent every police officer from observing the characteristics of suspects that may lead to discriminatory behavior, it is suggested that biases be eliminated (or at least reduced) in later stages of the decision-making process through evidence admissibility rules in the adjudication process.[99] An immediate implication from our theory of contagious discrimination is that such commendable efforts may be overly optimistic and insufficient, since discriminatory practices by some members of society can propagate to those who do not engage in the types of statistical or taste-based discrimination. Consequently, despite the efforts of well-intentioned individuals to act fairly and the development of institutional features aimed at reducing discrimination, the impact of discriminatory practices by some can be profound, potentially leading to discriminatory behavior by otherwise non-discriminatory decision makers, such as judges and prosecutors.

In Section III.A, we provide a brief and intuitive explanation of the dynamics that emerge in the model we present in Section III.B. In Section III.B, we first introduce our model and subsequently use it to formalize the concept of contagious discrimination. In Section III.C, we discuss how this concept relates to other theories of discrimination proposed in the literature.

A. Conception

Consider an often-reported form of discrimination: Everything else being equal, a Black citizen in the United States is more likely than a white person to be stopped and cited by the police.[100] To illustrate our theory of discrimination, we assume that while the upstream police may be discriminatory, the downstream legal system does not discriminate when rendering the final decision (acquittal or conviction). Nevertheless, we show

99 *See, e.g.*, Mungan, *supra* note 52, at 312–13 (reviewing the literature).

100 *See, e.g.*, Sonja B. Starr, *Explaining Race Gaps in Policing: Normative and Empirical Challenges* 5–6 (Univ. of Mich. L. & Econ. Working Papers, Art. 110, 2015), http://repository.law.umich.edu/law_econ_current/110 [https://perma.cc/T3A3-54ZJ]; Carvalho, Mizael & Sampaio, *supra* note 5.

that the Black suspect may still receive a worse sentence than a similarly situated white suspect through the plea-bargain process. In other words, *even if* the legal system aims to correct the problem of police discriminatory practices by treating all suspects fairly without regard to race, discriminatory practices by the police can still permeate into the judicial system.

The intuition behind this phenomenon is that people from different groups may rationally perceive the expected effect associated with a conviction record (stigmatization) differently. Even if a Black suspect is not convicted in the present period, she may believe that she is more likely to be cited again later and, therefore, more likely to be stigmatized in the future: She expects the police to be discriminatory, even if she expects the judicial system to be non-discriminatory.[101] The benefit of carrying a clean record relative to carrying a stigmatized record is less for the discriminated individual because she faces higher probabilities of *future* citation, conviction, and stigmatization.

All else equal, accepting the *same* plea-bargain is thus less appealing to a white individual[102] because acquittal and its associated clean record are more attractive to her than to a Black individual. Even though the prosecutor and the judicial system are assumed to be neutral and fair, the Black suspect may therefore still end up being discriminated against in the plea-bargain stage due to the higher probability of facing future charges. The next Section formalizes this result.

B. Model

We introduce a two-period model in which individuals interact with various decision-makers over two time periods.[103] In traditional economic analyses, these models are used to isolate the issue of interest, while simplifying other aspects that are not directly relevant to the analysis.[104] The two periods in these models are assumed to reflect the dynamic nature of the interactions. This is similarly a simplifying assumption and the insights obtained from the model are valid in more realistic models containing more periods as well.[105] Simplification allows us to focus on the most important

---

101 We assume judges do not discriminate to isolate the effect from contagious discrimination. If a suspect also expects downstream judges or juries to be biased, then she is even more likely to accept a worse plea-bargain.

102 *See, e.g.*, Kutateladze, Andiloro, Johnson & Spohn, *supra* note 96 (finding Black and Hispanic defendants are less likely than white defendants to receive a reduced charge offer from prosecutors, controlling for crime severity and criminal history).

103 The assumption of two time periods is for the purpose of simplification. The model can easily be extended to include more than two periods.

104 For example, see the overlapping generation model described in LARS LJUNGQVIST & THOMAS J. SARGENT, RECURSIVE MACROECONOMIC THEORY 331–78 (4th ed. 2018).

105 *See generally* ROBERT GIBBONS, GAME THEORY FOR APPLIED ECONOMISTS (1992) (detailing real-world applications of game-theoretic models).

factors and relationships without being bogged down by extraneous details.

In each time period, multiple decision makers sequentially make a decision that affects an individual who may be discriminated against. We assume upstream decision makers are discriminatory whereas downstream actors, who make the final decision, are not discriminatory.[106]

In our setting, there are three stages of the decision-making process:

The first stage of the decision-making process involves a police officer deciding whether to charge (or cite) a suspect. A suspect is cited with probability p^d, which is a number between 0% and 100%. The superscript d highlights the fact that this player (the police) may practice discrimination. For example, if race is the focus, then d can be w for white people or b for Black people with $p^w < p^b$ indicating that a Black person is more likely to be cited than a white person. Therefore, p^d captures the degree of discrimination: the higher the difference in the values of these probabilities (e.g., between p^w and p^b), the stronger the degree of discrimination.

The second-stage decision maker is the prosecutor who makes a plea-bargain offer to the suspect. If the offer is not accepted, the suspect goes to the third-stage decision maker—the judge—for adjudication. If the case is adjudicated by judges, the suspect is found guilty with probability q.

This model focuses on how discrimination in the first stage might affect the behavior of downstream actors in the legal system: the prosecutors and judges, who are assumed to engage in neither taste-based nor statistical discrimination. Therefore, there is no subscript on q, as we assume downstream judges do not discriminate and suspects correctly believe that.[107] We show that upstream discrimination by the police (modeled by the difference in value of p^d for different groups) can affect the downstream final disposition of the case.

Note also that both p^d and q are independent of the period number or conviction status: We assume the probabilities of receiving a police citation and conviction do not depend on whether an individual has a criminal record. This simplification allows us to isolate the contagious effect of discrimination that our theory highlights.

If a suspect is acquitted, there is no penalty. On the other hand, if a suspect is found guilty, she receives a penalty (e.g., imprisonment, a fine, or other sanction) which carries a cost of z and also suffers stigmatization costs s associated with a criminal record. The extralegal stigmatization costs s reflect the adverse behavior of the general public towards people with

106 We assume the upstream actor is discriminatory and the downstream actor is not because if the downstream, final-decision maker is discriminatory, then clearly the final decision will be discriminatory.

107 One may argue that since discriminated suspects are arrested with higher frequency due to discrimination, they are less likely to be convicted by the judge. The results of our study, however, will not change if judges cannot perfectly correct for this difference in input.

criminal records.[108] For example, people may shun ex-convicts,[109] and, especially in jurisdictions without ban-the-box laws, employers may prefer to hire candidates without a criminal record.[110]

We assume every individual has a clean record at the beginning of the first period. If one is not cited or is acquitted, that person enters the second period with a clean record: There will be no stigmatization cost s suffered by her in the first period, and in the second period she will suffer stigmatization cost s if and only if she is cited and found guilty, either through trial or plea, in that period. On the other hand, a cited individual who accepts a plea-bargain offer, or who rejects it and is subsequently convicted, suffers stigmatization cost s in the first period and suffers the same again in the second period.[111]

Notice the model does not assume whether the suspect is actually guilty or innocent. Our analysis compares the impact of discriminatory policing on an otherwise *similarly situated* suspect in the plea-bargaining process.[112] The same result of disparate impact would result whether we assume this similarly situated suspect is innocent or guilty.[113]

We present and analyze our model through backward induction by considering the second period before moving on to analyze the first period. This is a common method for analyzing multi-period interactions, in which individuals consider the future consequences of their actions while making

---

108 *See* Ericka B. Adams, Elsa Y. Chen & Sarah E. Lageson, *The Symbiotic Harm of a Criminal Record*, 53 CRIM. JUST. & BEHAV. 403, 404–05 (2026) (reviewing the negative collateral consequences of criminal records in terms of employment, housing, and other aspects of economic stability and security).

109 *Id.*

110 *See, e.g.*, Amanda Agan & Sonja Starr, *Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment*, 133 Q.J. ECON. 191, 194–95 (2018) (finding that individuals with criminal records have a 63 percent higher likelihood of callbacks in jurisdictions with ban-the-box laws than in jurisdictions without ban-the-box); Devah Pager, *The Mark of a Criminal Record*, 108 AM. J. SOCIO. 937, 960 (2003) (showing the negative impact of a criminal record on employment opportunities); Harry J. Holzer, Steven Raphael & Michael A. Stoll, *The Effect of an Applicant's Criminal History on Employer Hiring Decisions and Screening Practices: Evidence from Los Angeles*, *in* BARRIERS TO REENTRY? THE LABOR MARKET FOR RELEASED PRISONERS IN POST-INDUSTRIAL AMERICA 117, 122 (Shawn Bushway, Michael A. Stoll & David F. Weiman eds. 2007) (survey data demonstrating how employers' willingness to hire ex-offenders is very limited).

111 This implies that the second-period stigma is independent of the times the person has been convicted in the past. This is a simplifying assumption. In many settings, one can expect the stigma from the second offense to be less than the stigma from the first. *See* Murat C. Mungan, *Wrongful Convictions, Deterrence, and Stigma Dilution*, 25 SUP. CT. ECON. REV. 199, 211 (2018) (finding that stigma associated with having a criminal record diminishes with each conviction).

112 We are thus comparing an innocent (guilty) suspect who is discriminated against to a similarly positioned innocent (guilty) suspect who is not discriminated against.

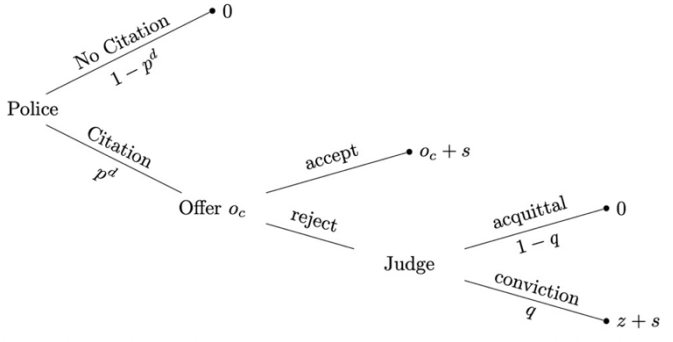
113 A potential extension complicates the model by allowing individuals to choose whether to commit a crime or not at the beginning of each period. Under this extension, whether the individual is discriminated against or not will affect that person's choice whether to commit a crime.

decisions in the present.[114]

1. Second-Period Analysis

Individuals enter the second period with either a clean record or with a criminal record. We begin our analysis with the individual entering the second period with a clean record. The figure below depicts the three stages in the second period for an individual with a clean record. At the end of each terminal node (the end of each path in the figure) is a variable that denotes the *cost* to the individual.[115]

FIGURE 2. AN INDIVIDUAL WITH A CLEAN RECORD (PERIOD 2)



At the beginning of the second period, an individual with a clean record is cited by the police with probability p^d. If the individual is not charged or cited, the cost to her is zero because there is neither a penalty nor stigmatization. For simplicity, we assume that receiving a citation in itself does not result in stigma—individuals are stigmatized only if they are convicted, either through a plea or through trial.[116]

If the individual is cited, the prosecutor makes a plea-bargain offer o_c. Accepting the plea-bargain offer is construed as a conviction and therefore carries with it the additional cost of stigmatization s.[117] Rejecting the plea-

114 *See e.g.*, GIBBONS, *supra* note 105, at 55–73 (applying backwards induction to multi-stage games).

115 Very often cost is expressed with a negative sign in front. Since all of our utilities are in terms of cost, we choose to present cost without a minus sign for simplicity.

116 *See* Claude Fluet & Murat C. Mungan, *Informational Properties of Liability Regimes*, 54 J. LEGAL STUD. 199 (2025) (discussing the validity of this simplifying assumption).

117 The existing literature discusses the validity of this assumption in different contexts. *See, e.g.*, Cindy R. Alexander & Jennifer Arlen, *Does Conviction Matter? The Reputational and Collateral Effects of Corporate Crime*, *in* RESEARCH HANDBOOK ON CORPORATE CRIME AND FINANCIAL MISDEALING 87, 90 (Jennifer Arlen ed., 2018) (highlighting reputational damage for corporations who plead guilty in criminal settlements); Murat C. Mungan, *Optimal Non-Prosecution Agreements and the Reputational Effects of Convictions*, 59 INT'L REV. L. & ECON.

bargain offer moves the case to trial, in which the cited suspect is convicted with probability q. The acquittal of a suspect with a clean record involves neither criminal penalty nor stigmatization; the conviction of a suspect with a clean record involves both penalty cost z and stigmatization cost s.

As is often done in the law and economics literature, we assume throughout the paper that the prosecutor makes the least favorable plea-bargain offer that is acceptable to the cited suspect who is risk-neutral as a benchmark.[118] It is worth noting that this is a purely simplifying assumption, and the analysis remains unchanged even if the bargain is resolved through a splitting rule that reflects the parties' bargaining powers as described in Section II.B. The less favorable the defendant's outside option, the worse the bargaining outcome for the defendant, all else being equal.[119]

The prosecutor's offer o_c is thus the expected trial outcome:

$$o_c + s = q(s + z) + (1 - q)(0)$$
$$\Rightarrow o_c = qz - (1 - q)s \quad (1)$$

The expression implies that the plea-bargain sentence offered by the prosecutor is the suspect's expected punishment qz reduced by the expected avoidance of stigmatization $(1 - q)s$. With this offer in place, the defendant is indifferent between accepting the offer and rejecting it to proceed to trial. To simplify the analysis, we assume that an indifferent person takes the plea

57, 57–58 (2019) (same).

118 This is equivalent to the assumption that prosecutors act to maximize convictions. *See, e.g.*, William M. Landes, *An Economic Analysis of the Courts*, 14 J.L. & ECON. 61, 98 (1971) (laying out the assumption that prosecutors maximize their expected number of convictions); Jennifer F. Reinganum, *Plea Bargaining and Prosecutorial Discretion*, 78 AM. ECON. REV. 713, 717, 719 (1988) (detailing assumed goals of prosecutors, namely "appropriate punishment of the guilty, avoidance of punishment of the innocent and the conservation of resources spent on trials"); Thomas J. Miceli, *Plea Bargaining and Deterrence: An Institutional Approach*, 3 EUR. J.L. & ECON. 249, 260 (1996) (stating that certain models assume prosecutors are concerned with maximizing expected punishments); Scott Baker & Claudio Mezzetti, *Prosecutorial Resources, Plea Bargaining, and the Decision to Go to Trial*, 17 J.L. ECON. & ORG. 149, 155 (2001) (citing literature that identifies conviction maximization as prosecutors' main goal). It can be argued that the prosecutor, by making a worse offer to the discriminated individual, is not treating the individual neutrally or fairly. However, the plea-bargain process is a settlement negotiation in which the BATNA (the best alternative to a negotiated agreement) of each side of the bargaining table makes an impact on the final negotiation outcome. The prosecutor therefore can be viewed as merely bargaining with the defendant or the defendant's lawyers to come to an agreement without any awareness of the defendant's characteristics such as race. Instead, it is the defendant's own expectation of his or her BATNA that ultimately leads to a worse outcome for the discriminated individual. More generally, we need not make the simplifying assumption that the bargaining outcome is the least favorable to the defendant—instead, what matters is the recognition that *ceteris paribus*, the discriminated individual's expectation of worse future treatment has an impact on the individual's BATNA, leading to a worse bargaining outcome in the present period.

119 *See infra* Section II.B (reviewing theoretical and empirical studies that support or refute the notion).

offer.[120]

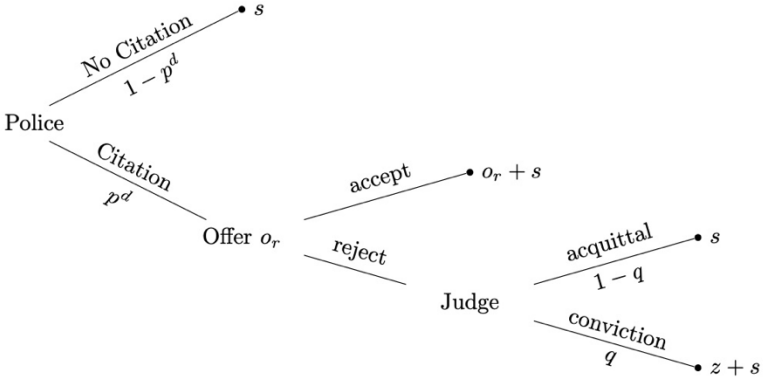
The expected cost for the individual with a clean record at the beginning of the second period is:

$$V_c = (1 - p^d)(0) + p^d\{q(s + z) + (1 - q)(0)\}$$
$$\Rightarrow V_c = p^d q(s + z) \quad (2)$$

There are two parts to this equation corresponding to two possibilities: (1) With probability $1 - p^d$ the person expects not to be charged by the police and therefore will face no cost since there is no prior stigmatization. (2) With probability p^d, the person expects to be charged by the police. The individual understands that fighting the case in court will result in no sentence with probability $1 - q$ and a conviction with probability q. A conviction in turn brings a cost of the sentence itself z and the cost of stigmatization s.

Next, we analyze the second period for a stigmatized individual, which is depicted in the figure below. A key difference is that the stigmatized individual always suffers stigmatization cost s in the second period no matter what transpires—we assume no double stigmatization.[121]

FIGURE 3. AN ALREADY STIGMATIZED INDIVIDUAL (PERIOD 2)



At the beginning of the second period, a stigmatized individual is cited by the police with probability p^d. Even if the individual is not cited, she still faces the stigmatization cost s. If the stigmatized individual is cited, the

120 It is standard in this type of modeling exercise to assume that an indifferent individual selects one option or the other. This assumption serves only as a simplifying device, since the defendant's expected costs under both options are identical.

121 *See supra* note 111 regarding this assumption.We conjecture that our results would remain the same so long as there are diminishing returns to double stigmatization. A micro-foundation for this assumption can be found in Mungan, *supra* note 111.

prosecutor makes a plea-bargain offer o_r, where the subscript r indicates the presence of a record. Accepting the plea-bargain offer results in a cost of $o_r + s$ since there is no double-stigmatization. Rejecting the plea-bargain offer moves the case to trial, in which case the cited suspect is convicted with probability q. The acquittal of the stigmatized suspect with probability $1 - q$ carries a stigmatization cost of s only; conviction with probability q results in both criminal penalty z and stigmatization cost s.

The prosecutor's offer o_r is the expected trial outcome:

$$o_r + s = q(s + z) + (1 - q)(s)$$
$$\Rightarrow o_r = qz$$

The stigmatized individual receives a plea-bargain offer equal to his expected punishment qz. Compared to the offer made to the suspect with a clean record, there is no reduction of the avoidance of stigmatization $(1 - q)s$ since the stigmatized suspect suffers reputational costs regardless of what happens in the second period.

The expected cost for the individual with a record at the beginning of the second period is:

$$V_r = \left(1 - p^d\right)(s) + p^d\{q(s + z) + (1 - q)(s)\}$$
$$\Rightarrow V_r = s + p^d qz. \quad (3)$$

Equation (3) expresses the cost an individual expects to suffer if she enters the second period with stigma.

We have thus far established that for the same cited suspect in the second period, a clean record makes her better off because the plea-bargain offered is shortened by $(1 - q)s$ for the individual with a clean record in the second period:

$$o_r - o_c = qz - qz + (1 - q)s = (1 - q)s > 0.$$

Similarly, the expected cost for the individual entering the second period with a criminal record is higher than that for the individual with a clean record:

$$V_r - V_c = p^d qz + s - p^d q(s + z) = \left(1 - p^d q\right)s > 0. \quad (4)$$

Equation (4) expresses the marginal cost of stigmatization in the second period.

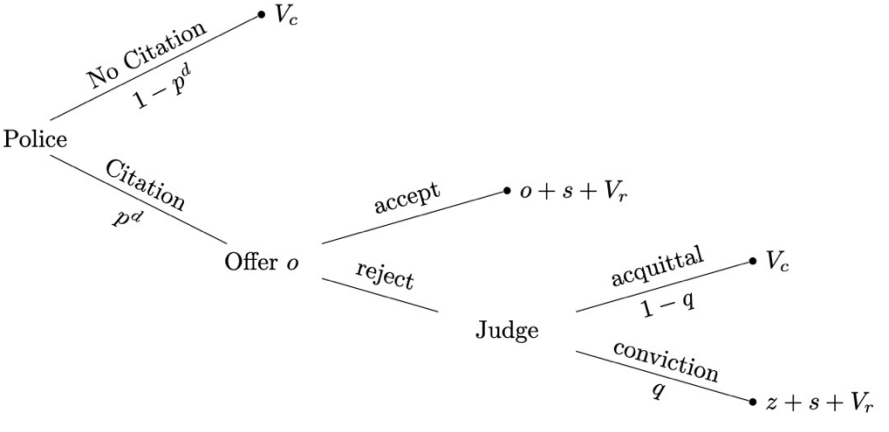
There are a few main takeaways from this. First, the expression is positive, indicating that stigmatization imposes additional costs on an

individual in the second period. Second, the marginal cost of stigmatization decreases in p^d: the more an individual is discriminated against, the lower the increase in marginal cost of stigmatization. The explanation for the decreasing rate of stigmatization in the degree of discrimination is that a discriminated individual who maintains a clean record in the first period is less likely to continue avoiding stigmatization in the second period than a non-discriminated person.

2. First-Period Analysis

Figure 4 below depicts the first-period interaction. All individuals enter the first period with a clean record. The only difference from the setting of an individual entering the *second* period with a clean record is the individual in the first period also considers the cost she will face in the second period, which depends on whether she enters the second period with a clean record (facing expected second-period cost V_c) or a stigmatized record (facing expected second-period cost V_r).

FIGURE 4. THREE STAGES FOR ALL INDIVIDUALS (PERIOD 1)



With probability p^d, an individual is cited by the police. If the individual is not cited, the first-period cost is 0 because there is neither a penalty nor stigmatization. Moreover, the individual faces expected cost V_c for entering the second period with a clean record. If the individual is cited, the prosecutor makes a plea-bargain offer o. Note that we use no subscript to indicate that this is a first-period offer since all individuals have a clean record at the beginning of the first period. Accepting the plea-bargain offer is construed as a conviction and therefore carries with it the additional cost of stigmatization s in the first period and the individual will enter the second period with expected cost V_r. Rejecting the plea-bargain offer moves the case

to trial, in which the cited suspect is convicted with probability q. The acquittal of a suspect with a clean record involves neither a penalty nor stigmatization, and the individual enters the second period with a clean record facing expected cost V_c. The conviction of a suspect, on the other hand, involves both penalty cost z and stigmatization cost s in the first period with the individual entering the second period with stigma facing expected cost V_r.

Due to reasons explained in the analysis of the second period, the prosecutor makes an offer that leaves the defendant indifferent between accepting and rejecting it. Thus, the offer o is characterized by the following equation:

$$o + s + V_r = q(s + z + V_r) + (1 - q)(0 + V_c)$$
$$\Rightarrow o = qz - (1 - q)s - (1 - q)(V_r - V_c)$$
$$\Rightarrow o = qz - (1 - q)s - (1 - q)\left(1 - p^d q\right)s$$

Comparing this first-period offer to second-period equilibrium offer $o_r = qz$ to the individual with stigma and $o_c = qz - (1 - q)s$ to the individual with a clean record, we see that $o < o_c < o_r$.

Moreover, whereas second-period plea-bargain offers are independent of whether the individual belongs to the discriminated group, the first-period plea-bargain offer o depends on whether the individual is discriminated against. Because o is increasing in p^d, the discriminated individual, facing a higher probability p^d of receiving a citation by the police, receives a longer plea-bargain sentence. Thus, the non-discriminatory prosecutor effectively makes a longer and worse first-period plea-bargain offer to the discriminated individual even though the prosecutor does not prefer penalizing the discriminated group and does not engage in the type of statistical discrimination we have previously reviewed.

This reveals our key result: Even if the judiciary is non-discriminatory and the prosecutor has no preference for discrimination, the downstream prosecutor ends up making a discriminatory plea-bargain offer to the discriminated individual. And this discriminatory outcome would not have occurred but for the discriminatory practices of the upstream police officers.

To complete our analysis, the expected cost at the beginning of the first period is computed as follows:

$$V = \left(1 - p^d\right)(0 + V_c) + p^d\{q(s + z + V_r) + (1 - q)(0 + V_c)\}$$
$$\Rightarrow V = V_c + p^d q(s + z + V_r - V_c)$$
$$\Rightarrow V = p^d q\left(2z + 3s - sp^d q\right)$$

This expression is increasing in p^d: the greater the probability of

receiving a citation by the police, the higher the expected cost the individual faces at the beginning of the first period.[122] As expected, a discriminated minority member is worse off at the beginning of the first period than a majority member.

C. Relationship to Other Theories of Discrimination

It may be worth exploring how the occurrence of discrimination at the plea stage differs from the types of discrimination previously discussed. Here, we briefly describe this in the specific context studied in our model.

Our theory is clearly different from taste-based discrimination because we do not assume the prosecutor has an intrinsic preference for individuals with certain characteristics.

Our theory is also different from statistical discrimination, in which the decision maker discriminates based on the suspect's characteristics in order to optimize under incomplete information. Under our theory, the discriminatory outcome occurs even if the decision maker (i.e., the prosecutor in our model) cannot observe the suspect's characteristics. Because the suspect has a worse outside option due to outside discrimination, the suspect has a worse bargaining position and therefore is more likely to end up with a worse plea-bargaining deal than a similarly positioned suspect who faces no or less outside discrimination. Under our theory, the prosecutor does not discriminate based on the suspect's characteristics but agrees to a worse plea-bargaining offer for the discriminated suspect who is more willing to accept the same plea offer than a suspect who does not face the same discrimination. The novelty of our theory thus lies not in how prosecutors make inferences based on race, but that such discrimination by the prosecutor would not have taken place but for the discriminatory practices by other institutional players (i.e.,the police).

Our theory of contagious discrimination is perhaps the closest to institutionalized discrimination under sociological theories in that we have identified an institutional practice that can perpetuate discrimination. However, our theory differs from existing sociological works in that contagious discrimination would not have occurred but for the discrimination of certain outside individuals (i.e., the police), whereas the institutional discrimination theories thus far identified by sociologists do not rely on discrimination at the individual level. In other words, while both our theory and theirs share the same symptom of disparate outcome at the institutional level, theirs stems from the historical development of the institution itself whereas the cause of our theory is individualized

[122] $\mathrm{V} = p^d q(2z + 3s - sp^d q)$ is increasing in p^d because its derivative with respect to p^d is $2zq + 3sq - 2sp^d q^2 = 2zq + sq + 2sq(1 - p^d q) \geq 2zq + sq \geq 0$ since the variables p^d and q are probabilities between 0 and 1. Notice that the derivative is strictly positive unless $q = 0$.

discrimination by other actors.

IV
DISCUSSION AND IMPLICATIONS FOR PLEA BARGAINING

An immediate implication of our theory is that society cannot solely focus on debiasing final-decision makers to eradicate discrimination. Targeting disparate outcomes at the prosecutorial and adjudicatory levels to identify discrimination and remedying it through policies aimed at prosecutors and judges may not be effective if the root cause of the problem lies in discrimination elsewhere.

Our theory underscores the need to examine the interactive nature of discrimination among different institutional players. In particular, we highlight a problematic aspect of interactive bargaining that has been overlooked in the literature. Consequently, our theory offers an additional critique of the plea-bargaining process within the U.S. criminal justice system.

It is important to note though that this assumption remains valid even if not all police officers engage in discriminatory behavior. A defendant could be cited by an officer who does not discriminate in the initial period.[123] It is the anticipation of future discrimination by another party that can lead to disparate impact in an otherwise discrimination-free environment. The expectation of being discriminated against by some police officers in the future prompts the defendant to accept a less favorable plea deal: The discriminated defendant anticipates a higher likelihood of returning to the system and thus faces a relatively lower cost of acquiring a criminal record in the current period.

A. Defendant Perceptions Matter

We have explicitly modeled the case in which some defendants face actual outside discriminatory practices. However, contagious discrimination can still arise even if the outside discrimination is merely an incorrect perception held by the discriminated defendants. If some individuals (incorrectly) believe that some police officers are more likely to arrest them due to discrimination, they will perceive less favorable outside options and be more inclined to accept worse plea-bargaining deals.

Thus, the first implication is that correcting upward misperceptions of anticipated discrimination can lead to reduced contagious discrimination.

123 One extension of our model is to assume that the defendant can face different types of police officers in period 1—each person encounters a police officer who may either be racist, or not, with some probability r. Thus, $1 - r$ proportion of all defendants in the first period have not yet been discriminated against. However, even among those defendants who have not been discriminated, their anticipation of being discriminated against in the future will change their outside options.

For example, suppose only ten percent of police officers discriminate against a group of defendants, but members of this group mistakenly believe that the probability of facing discrimination is ninety percent. This misperception might prompt these defendants to accept worse plea offers than they would if they held accurate beliefs.

While such divergence in belief could potentially be mitigated by increasing transparency in arrest data, a particularly challenging aspect of correcting misperceptions is confirmation bias. At the plea-bargaining stage, the defendant has just experienced an arrest, which may make it difficult for her to accept that the actual probability of an arrest is lower than she believes.[124] This is because the defendant is more likely to use the recent arrest as evidence to confirm her belief that police officers are discriminatory, even if the arresting officer did not actually discriminate.[125]

Therefore, it is crucial to enhance procedural transparency by requiring police officers to provide justifications for each arrest and allowing the suspect's lawyer to review these justifications with their clients. Additionally, the use of technology (such as body cameras[126]) and passing laws to set up committees to review police stop data (such as the California Racial and Identity Profiling Advisory Act (AB 953)) could further increase police accountability.

B. Contagious Discrimination Reinforces Statistical Discrimination

Under contagious discrimination, some suspects will agree to worse sentences due to their belief that they are more likely to be arrested in the future due to outside discrimination. This disparate outcome in turn reinforces the belief that members of the discriminated group are more likely to be found guilty and receive harsher sentences. Thus, if police officers based their arrest decisions on how likely it is that the arrestees will be convicted and receive criminal sanctions, contagious discrimination and statistical discrimination can mutually reinforce each other.

To mitigate this cycle of reinforcement, it is crucial to (1) limit the extent to which police officers rely on observed incarceration statistics in their arrest decisions or (2) ensure that police officers are not evaluated based on the accuracy or final judicial outcomes of the arrests they make.

---

124 *See* Michael D. Schlosser, Jennifer K. Robbennolt, Daniel M. Blumberg & Konstantinos Papazoglou, *Confirmation Bias: A Barrier to Community Policing*, 6 J. CMTY. SAFETY & WELL-BEING 162 (2021) (reviewing confirmation bias in policing).

125 Psychologists call this "confirmation bias." *See generally id.* (discussing confirmation bias in community policing).

126 *See* Anthony A. Braga, William H. Sousa, James R. Coldren, Jr. & Denise Rodriguez, *The Effects of Body-Worn Cameras on Police Activity and Police-Citizen Encounters: A Randomized Controlled Trial*, 108 J. CRIM. L. & CRIMINOLOGY 511, 535 (2018) (evaluating the effectiveness of body-worn cameras on police behavior).

C. Legal Representation and Plea Bargain Institution Reform

One of the key driving forces behind our findings is the interactive process of bargaining. All else being equal, the worse one's outside option, the more challenging it is for one to reach a favorable deal at the bargaining stage. The unfavorable outcome arises from the suspect's expectation of outside discrimination.

In our criminal justice system, criminal defendants are entitled to legal counsel.[127] In plea-bargaining, defense attorneys negotiate with prosecutors on behalf of their clients. Just as effective representation by the defense lawyer can significantly influence the plea-bargaining outcome,[128] lawyers can also provide an additional layer of protection for defendants against contagious discrimination.

A thought experiment may clarify this point. Before engaging in plea-bargaining on behalf of her client, a defense lawyer attempts to determine her client's outside option. Suppose her client indicates that a five-year sentence is the maximum she is willing to accept. The lawyer, drawing on case-related information and her experience, has strong reason to believe that a similarly positioned suspect of a different race would not agree to any sentence longer than three years.

Assuming these figures are accurate, the defendant presents a worse outside option due to her correct perception of outside discrimination. While it may seem beneficial for the lawyer to advise her client to disregard her race in making her assessment, such advice does not change the actual outside option that the client faces. So, what can lawyers do without violating their ethical duty to faithfully represent their clients?

One possibility is for lawyers and prosecutors to focus their negotiations on the merits of the case rather than on outside options. Judge Gerard Lynch has pointed out that prosecutors are public figures seeking justice, not the highest price.[129] Defense attorneys and prosecutors may be able to reduce contagious discrimination by adopting a more cooperative bargaining approach, instead of a competitive one that aims to split the surplus.

In fact, Judge Gerard Lynch has further suggested that plea-bargaining is not haggling in a bazaar, but resembles a standard process:

> Many, perhaps most, cases are processed pursuant to fairly standard rules: the policies of a particular United States Attorney may permit a drug

---

127 *See* U.S. CONST. amend. VI.

128 *See* Kelsey S. Henderson & Lora M. Levett, *Plea Bargaining: The Influence of Counsel*, *in* 4 ADVANCES IN PSYCHOLOGY AND LAW 73, 85–86 (Brian H. Bornstein & Monica K. Miller eds., 2019) (reviewing research demonstrating how attorneys influence the plea-bargaining process).

129 Gerard E. Lynch, *Our Administrative System of Criminal Justice*, 66 FORDHAM L. REV. 2117, 2131 (1998).

> courier or "mule" to plead to a standard reduced charge, for example, or a district attorney may offer any mugger who does no physical injury a standard disposition (perhaps further varied depending on whether a weapon was used) somewhere short of the theoretically-applicable offense of robbery. The rules are more like those of the supermarket than those of the flea market: there is a fixed price tag on the case, and you will get no farther "bargaining" with the prosecutor than you will by making a counteroffer on the price of a can of beans at the grocery.[130]

This description, however, has been strongly and persuasively challenged by Professors Ronald Wright and Marc Miller.[131] They argue that the current system lacks transparency and provides prosecutors with too much flexibility, resulting in "too much opportunity to treat similarly situated defendants differently."[132] Without delving into this important debate, we note that the ideal described by Judge Gerard Lynch—standard rules for sentence reduction in plea-bargaining—could be a viable solution to reduce contagious discrimination.

CONCLUSION

Given the pervasive harms of discrimination, developing theories to understand its causes and exacerbating factors is an important first step in designing policies to eliminate it. Existing theories have been highly influential, with much thought going into testing and identifying the different forms of discrimination in specific settings.[133] None of the theories, to our knowledge, examines the interaction of different institutional players. This Essay introduces a novel theory to fill this gap. In a setting of repeated interactions, our theory demonstrates that outside discrimination can creep into an otherwise discrimination-free decision process.

An important implication of our theory is that when there are multiple interacting parts within a system in which discrimination can take place, targeting a particular level of the system to reduce discrimination may be insufficient. Just as medical treatment often targets the cause of a disease

---

130 *Id.* at 2130.

131 Ronald Wright & Marc Miller, *Honesty and Opacity in Charge Bargains*, 55 STAN. L. REV. 1409, 1409–10 (2003).

132 *Id.* at 1411.

133 *See generally* Altonji & Pierret, *supra* note 55 (employment discrimination based on education); Knowles, Persico & Todd, *supra* note 26 (discrimination in motor vehicle searches); Kerwin K. Charles & Jonathan Guryan, *Prejudice and Wages: An Empirical Assessment of Becker's* The Economics of Discrimination, 116 J. POL. ECON. 773 (2008) (racial bias in wages); Steven D. Levitt, *Testing Theories of Discrimination: Evidence from* Weakest Link, 47 J.L. & ECON. 431 (2004) (discrimination in contestant voting behavior on a game show); John A. List, *The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field*, 119 Q.J. ECON. 49 (2004) (bias in the sports card market); Margaret A. Meyer & Richard Gonzalez, *Detecting Bias in Traffic Searches: Examining False Searches of Innocent Drivers*, 40 J. QUANT. CRIMINOLOGY 791 (2024) (racial bias in traffic searches).

instead of or in addition to its symptoms, our theory of contagious discrimination suggests that identifying the cause of discrimination and reducing it at its source may sometimes provide greater social returns.