

Articles

RACIAL DISCRIMINATION IN RETAILERS’ WILLINGNESS TO ACCEPT RETURNS: A FIELD STUDY

Meirav Furth-Matzkin

ABSTRACT—Black Americans have long faced discriminatory treatment while shopping in retail establishments, including, most notably, being subjected to increased surveillance, inconsistent pricing, and inferior customer service. Little attention, however, has been paid to other post-purchase aspects of retail transactions. Specifically, do Black Americans receive the same treatment as white customers when it comes to performing sellers’ formal policies or contracts? While it is understood that salespeople are often given discretion to deviate from standard form contracts, sometimes departing from the literal terms to satisfy consumers, there has been a notable absence of systematic exploration into how salespeople exercise this authority and whether racial disparities manifest in the exercise of such discretion. This Article reports on an original audit study that involved sending testers who were matched on all observable characteristics other than gender and race into downtown Chicago stores to return prepurchased items. All of the audited stores had official policies that required a receipt as a prerequisite to a return, but the testers did not present receipts for the items.

The findings revealed stark racial disparities: Black shoppers, both women and men, encountered more frequent denials and were less likely to be offered refunds than their white counterparts, with white women receiving the most favorable treatment. These disparities persisted and became even more pronounced after the testers asked to speak to management. Insights gleaned from novel interviews with store clerks suggest that these discrepancies stem from ingrained biases that associate race with social status and expected behavior.

In response, this Article outlines several recommendations that would reform the legal landscape to more robustly combat retail race discrimination in contractual performance. These include broadening the reach of existing legislation to explicitly outlaw racially discriminatory practices in the

execution of seller contracts and incentivizing sellers to implement strategies that have been shown to reduce bias in other decision-making environments.

AUTHOR—Assistant Professor, Tel Aviv University School of Law. This research was generously supported by the Coase-Sandor Institute at the University of Chicago. I thank Oren Bar-Gill, Omri Ben-Shahar, Adam Chilton, Alma Cohen, Hanoch Dagan, William Hubbard, Louis Kaplow, Tamar Kricheli-Katz, Florencia Marotta-Wurgler, Erin Miller, and Steven Shavell for helpful thoughts and suggestions. I also wish to thank the participants at the Annual Law & Economics Association Conference at Columbia University, the Berkeley Consumer Workshop, the Harvard Law & Economics Workshop, the Hebrew University Law & Economics Workshop, the NYU Law & Economics Workshop, the Stanford Law & Economics Workshop, the Tel Aviv Law Faculty Workshop, the University of Chicago’s Junior Scholars Forum in Law and Social Science, the Consumer Scholars Conference in Boston University, the Conference of Empirical Legal Studies (CELS) in Toronto, the Experimental Methods in Law Forum in Bonn, and the UCLA Junior Scholars Forum. The methods used in this study have been approved by the University of Chicago Institutional Review Board prior to data collection (IRB Study Number: IRB 18–1529). The study was preregistered in AsPredicted (see: “Post-Contract Discrimination in the Retail Market,” #16928, created 11/23/2018). Alona Zafrir, Laura Ash Smith, Nuri Albayrak, Raz Halfon, Einat Keinan, Daniel Malka, Roi Balsbalg, and Rhiannon Jones provided excellent research assistance.

INTRODUCTION	1137
I. BACKGROUND	1142
A. <i>"Shopping While Black": Past and Present</i>	1142
B. <i>Do Sellers Selectively Enforce "Standardized" Contracts or Policies?</i>	1146
II. THE CASE STUDY AND METHODOLOGICAL APPROACH	1150
A. <i>The Case Study: Receipt-Only Retail Returns</i>	1150
B. <i>The Methodological Approach</i>	1151
III. THE AUDIT STUDY	1155
A. <i>The Sample</i>	1155
B. <i>The Design of the Audits</i>	1156
C. <i>Results</i>	1159
D. <i>Discussion</i>	1177
E. <i>Implications and Generalizability</i>	1179
IV. POTENTIAL EXPLANATIONS	1181
A. <i>Psychological Research on Implicit Bias</i>	1181
B. <i>Insights Gleaned from Interviews with Store Clerks</i>	1186
V. POLICY RESPONSES	1190
A. <i>Prohibiting Discrimination in Contractual Performance</i>	1193
B. <i>Expanding Public Accommodation Laws</i>	1193
C. <i>Encouraging Sellers to Adopt Bias-Reducing Measures</i>	1194
CONCLUSION	1197
ONLINE APPENDIX	1199

INTRODUCTION

Federal law strictly prohibits racial discrimination in “making” and “enforcing” contracts as well as in places of public accommodation,¹ so—in the wake of the “Living While Black” and “Black Lives Matter” movements—one might assume racial discrimination in retail settings is a relic of the past.² But scholars have long documented the various problems

¹ Civil Rights Act of 1866, 42 U.S.C. § 1981 (“All persons . . . shall have the same right . . . to make and enforce contracts . . . as is enjoyed by white citizens . . .”). Section 1981(b) defines the term “make and enforce contracts” as “the making, performance, modification, and termination of contracts, and the enjoyment of all benefits, privileges, terms, and conditions of the contractual relationship.” *Id.*

² See, e.g., Taja-Nia Y. Henderson & Jamila Jefferson-Jones, *#LivingWhileBlack: Blackness as Nuisance*, 69 AM. U. L. REV. 863, 867–68 (2020) (listing the various viral incidents in which white people called the police on Black people that ultimately led to the hashtag’s inception). Indeed, this has been the

Black customers often face in retail establishments, a phenomenon known as “shopping while Black.”³ This term encompasses a range of discriminatory experiences in retail environments, including increased surveillance, inconsistent pricing, and generally inferior customer service.⁴

Instances of racial discrimination reported by Black consumers in public spaces, especially in retail settings, often involve overt experiences such as close monitoring by security guards or rude treatment from store agents.⁵ Yet there is a notable gap in evidence-based research concerning the unequal application of sellers’ standardized contracts or policies.

In particular, do Black customers receive the same treatment as their white counterparts when it comes to the enforcement of sellers’ seemingly uniform agreements or policies? A recent incident at a Walmart store in Texas suggests that there may be cause for concern. In that case, two Black men, one a former police officer and the other a pastor, faced accusations of theft when they attempted to return a defective television they had purchased earlier that day.⁶ In fact, the store summoned the police and the men were handcuffed, detained, and subjected to unwarranted scrutiny.⁷ Although the men were ultimately released due to a lack of evidence, the store agents denied them a refund and compelled them to sign a document pledging not to return to the store—a clear deviation from the store’s standard return policy.⁸ This incident, especially when coupled with anecdotal evidence of other Black customers’ experiences,⁹ raises the question as to whether there is a systemic pattern of differential treatment toward Black customers in contractual enforcement.

position expressed by courts as early as the 1960s. *See Newman v. Piggie Park Enters., Inc.*, 377 F.2d 433, 436 (4th Cir. 1967) (“Retail stores, food markets, and the like were excluded from [Title II of the 1964 Civil Rights Act] for the policy reason [that] there was little, if any, discrimination in the operation of them.” (citing 110 CONG. REC. 6533 (1954) (statement of Sen. Humphrey))), *aff’d*, 390 U.S. 400 (1968).

³ *See* Henderson & Jefferson-Jones, *supra* note 2, at 890 nn.143–44 and accompanying text.

⁴ *Id.*

⁵ These experiences of racial mistreatment and discrimination in retail environments are often collectively termed “shopping while Black.” For an overview of the “shopping while Black” phenomenon, see *infra* Section I.A.

⁶ Will Feuer, *Two Black Men Sue Walmart over Racial Discrimination, False Arrest*, N.Y. POST (Aug. 27, 2021), <https://nypost.com/2021/08/27/black-men-sue-walmart-claim-they-were-handcuffed-after-trying-to-return-defective-tv/> [<https://perma.cc/Y2TX-9KUS>]; Khristopher J. Brooks, *Two Black Men Say They Were Handcuffed While Returning a Broken TV to Walmart*, CBS NEWS (Aug. 27, 2021), <https://www.cbsnews.com/news/walmart-sued-houston-television-terence-richardson-dennis-stewart-lawsuit/> [<https://perma.cc/VH43-UP4S>].

⁷ Feuer, *supra* note 6; Brooks, *supra* note 6.

⁸ Feuer, *supra* note 6; Brooks, *supra* note 6.

⁹ *See infra* notes 33–36 and accompanying text.

Answering that question, however, is challenging because the difficulty of isolating race from intertwined factors—including one's personal and physical characteristics, socioeconomic background, personal history, and education, among other factors—complicates efforts to link a customer's race with a seller's conduct.¹⁰ This problem pervades race discrimination research.¹¹ Even with data correlating seller behavior with consumer demographics, establishing causality remains a challenge, as traditional regression analysis methods frequently fail to provide conclusive evidence of racial bias.¹²

To navigate these challenges, this Article introduces an innovative methodological framework. By combining comprehensive field audits and in-depth interviews with store clerks—the primary decision-makers in retail settings—this Article sheds light on the existence and prevalence of race-based discrimination in the enforcement of retailers' return policies.

Audit methodologies have traditionally been instrumental in investigating race and gender discrimination in various sectors, most predominantly in housing and employment.¹³ These audits usually involve testers who are meticulously matched on all relevant observable characteristics (such as age, attire, and education), other than gender and race. In the past, these studies have primarily focused on outcomes such as differential pricing or outright refusals to engage in transactions (such as hiring a job applicant, selling a product, or renting an apartment).¹⁴

¹⁰ See, e.g., Maya Sen & Omar Wasow, *Race as a Bundle of Sticks: Designs That Estimate Effects of Seemingly Immutable Characteristics*, 19 ANN. REV. POL. SCI. 499, 500 (2016) (proposing that race should be seen as a “bundle of sticks” and operationalized as such in experimental or empirical studies).

¹¹ See, e.g., Marianne Bertrand & Esther Duflo, *Field Experiments: Correspondence Studies*, in HANDBOOK ON ECONOMICS OF DISCRIMINATION AND AFFIRMATIVE ACTION 9, 152 (Ashwini Deshpande ed., 2023).

¹² A regression analysis attempts to reveal relationships between explanatory variables (for example, gender and race) and a dependent (or outcome) variable (for example, salary or hiring outcomes). Explanatory variables are the expected influences on the dependent variable. See, e.g., Daniel L. Rubinfeld, *Reference Guide on Multiple Regression*, in REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 305 (3d ed. 2011). For an overview of the limitations of regression analyses and the specific challenges in proving causation, see Bertrand & Duflo, *supra* note 11, at 152.

¹³ See Bertrand & Duflo, *supra* note 11, at 158–59, 169–70. In the consumer context, there have been relatively few attempts to study discrimination using audits. For a notable exception, see Ian Ayres, *Fair Driving: Gender and Race Discrimination in Retail Car Negotiations*, 104 HARV. L. REV. 817, 829 (1991), which found, based on an audit study, that car dealers frequently charged significantly higher prices for identical cars when dealing with Black customers and white female customers, as compared to similarly situated white male customers; and Ian Ayres & Peter Siegelman, *Race and Gender Discrimination in Bargaining for a New Car*, 85 AM. ECON. REV. 304, 305–09, 319 (1995), which replicated the car dealership discrimination study using a larger sample and obtained similar results.

¹⁴ For an elaborate survey of this literature, see generally Bertrand & Duflo, *supra* note 11.

This Article represents a pioneering effort to extend the use of the audit methodology to examine racial discrimination in the realm of contractual performance. By examining retail product returns, this Article aims to ascertain whether an optimistic view of racial progress in the retail setting can withstand empirical scrutiny. Specifically, this study was designed to focus on retail product returns, an area where Black customers have historically faced significant challenges.¹⁵

To test for racial discrimination in the context of retail product returns, I assembled and trained a group of 19 testers, comprised of 5 Black men, 4 Black women, 5 white men, and 5 white women—all of whom were undergraduate students residing in Chicago. To help ensure a consistent appearance, the testers dressed in a similar fashion. Their mission involved returning unworn clothing items, each of which had been prepurchased by other team members, to 59 retail stores in Chicago.¹⁶ Each store explicitly mandated a receipt for a return. Following a standardized script, the testers attempted to return the unused items, all in their original packaging, without receipts. In instances where store clerks did not offer an immediate refund, testers were instructed to request to speak with a manager. This structured approach enabled systematic examination of potential differences in treatment based on testers' race at different stages of the encounter.

The study shows that discrimination occurs not just overtly—such as outright refusal of service—but also in more subtle forms during post-contract enforcement. Significant racial disparities in treatment emerged: Black testers attempting to return items without receipts were about 8 percentage points less likely to obtain refunds than white testers under similar conditions, and about 14.5 percentage points *more* likely to experience complete denials of their return requests. These racial disparities not only persisted but also widened when testers voiced complaints or sought managerial intervention—Black testers were 25.2 percentage points more likely than comparable white testers to have their return requests fully denied.

The insights gathered from the interviews conducted with store clerks as part of this research suggest that the observed disparities in treatment

¹⁵ See *infra* Section I.A.

¹⁶ Chicago was chosen because it is a mixed city with a relatively large Black population. According to U.S. Census Bureau data, Chicago is almost 30% Black. *QuickFacts: Chicago City, Illinois; United States*, U.S. CENSUS BUREAU, <https://www.census.gov/quickfacts/fact/table/chicagocityillinois,US/PST045224> [https://perma.cc/PZY9-TARQ]. For a description of the stores' selection criteria and elaborate details on the design of the audit study and the data collection process, see *infra* Sections III.A–III.B.

may stem, at least in part, from implicit biases linking consumers' race to their perceived buying power, transactional expectations, and likelihood of lodging complaints.¹⁷ Previous research indicates that a heightened level of suspicion towards Black customers could also contribute to these biases.¹⁸

These findings paint a stark reality of ongoing racial inequality in retail environments, challenging the notion that race-based discrimination in shopping is a relic of the past. The discriminatory application of retail return policies, and possibly other standard consumer contracts, represents a largely unacknowledged form of everyday racial discrimination with significant adverse effects. This discrimination in contractual performance not only denies equal service but effectively constitutes a form of price discrimination: Black customers may be paying the same price, but they are receiving inferior service—a profound and troubling finding. In essence, Black customers seem to be subsidizing the shopping experience of their white counterparts.

The current legal framework falls short in adequately safeguarding consumers against discrimination in the execution of sellers' contracts. Although federal civil rights statutes, most notably 42 U.S.C. § 1981, explicitly prohibit racial discrimination in the “mak[ing] and enforc[ing]” of contracts, judicial interpretations have typically confined the statute's scope to the initial stages of contract formation.¹⁹ Recent legal developments hint at a possible expansion of this interpretation to include discrimination in post-purchase phases.²⁰ However, it remains uncertain whether courts will recognize practices such as granting disproportionate privileges or waiving contractual obligations based on race as forms of unlawful discrimination in contract enforcement.²¹

Building on the audit study's findings, this Article suggests that courts should consider embracing such a broader application of antidiscrimination laws.²² It also proposes that sellers consider implementing various bias-reducing measures, such as debiasing training, accountability mechanisms,

¹⁷ *Infra* Section IV.B.

¹⁸ *E.g.*, SHAUN L. GABBIDON & GEORGE E. HIGGINS, SHOPPING WHILE BLACK: CONSUMER RACIAL PROFILING IN AMERICA 10–19 (2020); George E. Schreer, Sandra Smith & Kirsten Thomas, “*Shopping While Black*”: Examining Racial Discrimination in a Retail Setting, 39 J. APPLIED SOC. PSYCH. 1432, 1432–33 (2009).

¹⁹ *See infra* Section V.A.

²⁰ *Infra* Section V.A.

²¹ *Infra* Section V.A.

²² For a more detailed examination of the proposal to prohibit sellers from discrimination in the performance of their form contracts, see Meirav Furth-Matzkin, *Banning Contractual Performance Discrimination*, 43 YALE L. & POL'Y REV. (forthcoming 2025).

and more diverse personnel.²³ These combined measures could hopefully mitigate biased discretionary practices in the implementation of sellers' formal policies and contracts.

This Article proceeds as follows. Part I contextualizes this study by situating it within existing literature on retail race discrimination and delving into the theoretical and empirical works that scrutinize the practical implementation of form contracts by sellers. Part II presents the case study and methodological approach. Part III provides an in-depth overview of the audit study, including details on the research design, the sample, and data collection methods. It then presents the results, which reveal significant racial disparities in the enforcement of retailers' return policies. Part IV explores potential explanations for the observed disparities, drawing on the insights gathered from in-depth interviews conducted with store clerks. Part V discusses potential policy implications in view of these findings. A brief conclusion follows.

I. BACKGROUND

A. *"Shopping While Black": Past and Present*

*"Name a store, any store, from Fifth Avenue to Main Street, and I'll bet I can find a Black person who has experienced discrimination there."*²⁴

Contemporary experiences of mistreatment in a retail setting are so commonly shared among Black consumers that there is even a term for it: "shopping while Black."²⁵ In the early 1990s, Texas University sociology professor Joseph Feagin documented some Black shoppers' experiences in retail stores by conducting in-depth interviews with thirty-seven Black, middle-class residents living in several American cities.²⁶ Based on those

²³ For a more in-depth exploration of how these bias-reducing measures could be implemented to reduce discrimination in contractual performance, see Meirav Furth-Matzkin, *Discrimination in Contractual Performance: Evidence, Theory, and Preliminary Policy Prescriptions*, 99 WASH. L. REV. 1165, 1185–90 (2024).

²⁴ Cassi Pittman Claytor, *"Shopping While Black": Yes, Bias Against Black Customers Is Real*, GUARDIAN (June 24, 2019), <https://www.theguardian.com/commentisfree/2019/jun/24/shopping-while-black-yes-bias-against-black-customers-is-real> [https://perma.cc/C24M-2D25].

²⁵ See *infra* note 36. Discrimination while shopping is most commonly associated with Black consumers, perhaps due to media coverage and highly publicized litigation involving Black consumers. Yet, other racial and ethnic minorities (including Hispanics, Asians, and Native Americans) report similar experiences of discrimination, also known as "shopping while brown or non-white." See, e.g., MICHELLE DUNLAP, *RETAIL RACISM: SHOPPING WHILE BLACK AND BROWN IN AMERICA* (2021) (compiling the stories of a wide variety of racial and ethnic minorities on their experience with discrimination in retail).

²⁶ Joe R. Feagin, *The Continuing Significance of Race: Antiblack Discrimination in Public Places*, 56 AM. SOCIO. REV. 101, 103 (1994).

interviews, Feagin concluded that, even while shopping, “a Black person cannot escape the stigma of being Black.”²⁷

This widespread discrimination has drawn attention from legal experts, civil rights advocates, and scholars, highlighting the ubiquity of such experiences among Black Americans.²⁸ As University of Pennsylvania Law Professor Regina Austin has observed, “[t]here can hardly be a [B]lack person in urban America who has not been denied entry to a store, closely watched, snubbed, questioned about her or his ability to pay for an item, or stopped and detained for shoplifting.”²⁹

In her memoir, *Inside Toyland: Working, Shopping, and Social Inequality*, Texas University sociology professor Christine Williams recounts her experience working as a salesclerk in two toy stores in the early 2000s, revealing jarring insights into the social and racial inequalities that permeated the toy-shopping experience at the time.³⁰ According to Williams, white customers, particularly white women, typically received treatment above and beyond store policies and expectations, while Black customers were rarely afforded the same consideration, even when they had defective merchandise or other legitimate reasons to complain.³¹

Williams illustrates her point using two specific incidents from her time as a clerk. The first involved an attempt by a white woman to purchase a Barbie doll from the *Wizard of Oz* collection using a “half-price coupon” that explicitly excluded the item.³² Despite the coupon’s explicit terms, the woman argued with two clerks until a manager was summoned and she was ultimately granted the discount.³³ In contrast, when a Black man tried to return a toy without a receipt, the store agents not only denied his request, but also summoned the police.³⁴ As Williams astutely observes, generally speaking, “a shouting white woman got whatever she wanted and a shouting [B]lack man got threatened with arrest.”³⁵

²⁷ *Id.* at 106–07.

²⁸ *See infra* notes 39–42.

²⁹ Regina Austin, “*A Nation of Thieves*”: Securing Black People’s Right to Shop and to Sell in White America, 1994 UTAH L. REV. 147, 148.

³⁰ CHRISTINE L. WILLIAMS, *INSIDE TOYLAND: WORKING, SHOPPING, AND SOCIAL INEQUALITY* (2006).

³¹ *See id.* at 130–32.

³² *Id.* at 112.

³³ *Id.*

³⁴ *Id.* at 127–28.

³⁵ *Id.* at 131.

These accounts are part of a broader, persistent pattern of racial bias and discrimination in public places, particularly in retail spaces.³⁶ Gallup polls conducted since 1997 consistently show that Black Americans report unfair treatment in retail environments more frequently than in any other setting.³⁷ As recently as July 2021, 35% of Black respondents reported encountering discrimination while shopping within the past thirty days.³⁸ Discriminatory experiences include being followed by security,³⁹ treated

³⁶ See generally GABBIDON & HIGGINS, *supra* note 18 (surveying available data sources, the history of consumer racial profiling, and criminological literature to show how Black shoppers are singled out for observation and discriminatory treatment by store agents); Cassi Pittman, “*Shopping While Black*”: *Black Consumers’ Management of Racial Stigma and Racial Profiling in Retail Settings*, 20 J. CONSUMER CULTURE 3 (2020) (drawing on qualitative interview data to examine Black shoppers’ experiences of retail racism and the cultural strategies adopted in response); Aronté Marie Bennett, Ronald Paul Hill & Kara Daddario, *Shopping While Nonwhite: Racial Discrimination Among Minority Consumers*, 49 J. CONSUMER AFF. 328 (2015) (reporting, based on a survey, that nonwhite customers are significantly more likely to feel discriminated against in stores than are white customers); David Crockett, Sonya A. Grier & Jacqueline A. Williams, *Coping with Marketplace Discrimination: An Exploration of the Experiences of Black Men*, 2003 ACAD. MKTG. SCI. REV., no. 4, at 1 (presenting results of in-depth interviews with Black males to illustrate race-related strain from marketplace discrimination and resulting coping strategies); IAN AYRES, *PERVASIVE PREJUDICE?: UNCONVENTIONAL EVIDENCE OF RACE AND GENDER DISCRIMINATION* 7–8 (2001) (suggesting that “[r]etailers may also discriminate in their willingness to accommodate private and somewhat idiosyncratic consumer requests” and citing research to that effect); Jennifer Lee, *The Salience of Race in Everyday Life: Black Customers’ Shopping Experiences in Black and White Neighborhoods*, 27 WORK & OCCUPATIONS 353 (2000) (reporting, based on seventy-five in-depth interviews of Black consumers, that Black customers feel that they are treated unfairly in shops located in predominantly white neighborhoods); JOE R. FEAGIN & MELVIN P. SIKES, *LIVING WITH RACISM: THE BLACK MIDDLE-CLASS EXPERIENCE* (1994) (using in-depth interviews with middle-class Black Americans to expose the persistent discrimination experienced in public places).

³⁷ Jeffrey M. Jones & Camille Lloyd, *Black Americans’ Reports of Mistreatment Steady or Higher*, GALLUP (July 27, 2021), <https://news.gallup.com/poll/352580/black-americans-reports-mistreatment-steady-higher.aspx> [<https://perma.cc/L8RD-ZUFU>].

³⁸ *Id.*

³⁹ See GABBIDON & HIGGINS, *supra* note 18, at 10–19 (discussing evidence that store agents tend to be more suspicious of Black customers, particularly Black men, and unfairly target them with surveillance and calls to the police).

rudely by clerks,⁴⁰ quoted higher prices for the same products,⁴¹ and denied service altogether.⁴²

Evidence of retail race discrimination has also received national attention due to the willingness of prominent Black Americans (including TV host Oprah Winfrey, basketball star Sheryl Swoopes, and Representative Maxine Waters) to share their personal stories of mistreatment and bias.⁴³ In her widely read memoir, Columbia Law Professor Patricia Williams described her firsthand experience with discrimination in a retail setting.⁴⁴ In one incident, she recalled being denied entry to a retail store by a white store agent while white customers were shopping inside the store for Christmas presents.⁴⁵ Williams famously described the “blizzard of rage” she felt when excluded from the store on account of her race.⁴⁶

Empirical studies further validate these personal accounts. Research shows that Black customers often receive inferior service, evidenced by longer wait times and heightened surveillance compared to their white

⁴⁰ See, e.g., JOSEPHINE LOUIE, HARV. UNIV., THE C.R. PROJECT, WE DON'T FEEL WELCOME HERE: AFRICAN AMERICANS AND HISPANICS IN METRO BOSTON ii (last revised April 25, 2005), <https://civilrightsproject.ucla.edu/research/metro-and-regional-inequalities/metro-boston-equity-initiative-1/we-don2019t-feel-welcome-here-african-americans-and-hispanics-in-metro-boston/louie-we-dont-feel-welcome-2005.0.pdf> [<https://perma.cc/J9S5-XN7P>] (finding, based on a survey, that over half of Black respondents residing in the Boston Metropolitan Area report being “treated with less respect, offered worse service, called names or insulted, or confronted with another form of day-to-day discrimination at least a few times a month”); Zachary W. Brewster & Sarah N. Rusche, *Quantitative Evidence of the Continuing Significance of Race: Tableside Racism in Full-Service Restaurants*, 43 J. BLACK STUD. 359, 375 (2012) (reporting on a study in which almost 40% of the surveyed restaurant workers reported providing lower quality service to minority consumers based on race).

⁴¹ See, e.g., Ayres, *supra* note 13, at 819 (finding that car dealers charged significantly higher prices for identical cars when dealing with Black customers, as compared to similarly situated white customers); Ayres & Siegelman, *supra* note 13, at 304–21 (replicating the dealership study using a larger sample and obtaining similar results).

⁴² See, e.g., GABBIDON & HIGGINS, *supra* note 18, at 14 (recounting cases from the Denny's restaurant chain where several high-profile Black patrons, including Secret Service agents and a federal judge, were denied service); Benjamin Edelman, Michael Luca & Dan Svirsky, *Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment*, AM. ECON. J.: APPLIED ECON., Apr. 2017, at 1, 2 (finding, based on an online field experiment, that Airbnb hosts were more likely to refuse booking requests made by guests with Black-sounding names compared to guests with white-sounding names); see also Andrew Hanson, Zackary Hawley, Hal Martin & Bo Liu, *Discrimination in Mortgage Lending: Evidence from a Correspondence Experiment*, 92 J. URB. ECON. 48, 48–49 (2016) (using similar methods and finding evidence of racial discrimination in the credit market).

⁴³ GABBIDON & HIGGINS, *supra* note 18, at 13.

⁴⁴ PATRICIA WILLIAMS, THE ALCHEMY OF RACE AND RIGHTS 44–45 (1991).

⁴⁵ *Id.*

⁴⁶ *Id.* at 45.

counterparts.⁴⁷ For instance, an observational study conducted in 2006 revealed that security guards disproportionately targeted Black consumers.⁴⁸ Similarly, a 2009 audit study observed increased scrutiny of Black customers by salespeople in high-end retail settings.⁴⁹

B. Do Sellers Selectively Enforce “Standardized” Contracts or Policies?

While it seems evident, both from Black respondents’ reports and the empirical research, that Black customers still suffer from mistreatment and discrimination in retail settings, the question remains, however, whether racial disparities exist in situations where consumers enter transactions that are purportedly governed by an explicit set of standard terms and conditions. Put differently, should we expect sellers to apply their form contracts differently based on race? The answer to this question has significant implications for our understanding of racial discrimination in consumer markets and the measures needed to address it effectively. Yet, despite its importance, no empirical research on the issue exists.

The paucity of research on this subject likely stems, at least in part, from the absence of an easily accessible database of relevant information.⁵⁰ As commentators have noted, obtaining “reliable data to confirm the regularity of consumer discrimination” remains challenging,⁵¹ especially within the context of mistreatment that occurs in retail spaces.⁵²

Furthermore, the specific form of discrimination under consideration—discrimination in the performance of a seller’s form contract or policy—may have contributed to the lack of research, as scholars have generally assumed that terms in form contracts are uniformly applied and that sellers

⁴⁷ Thomas L. Ainscough & Carol M. Motley, *Will You Help Me Please? The Effects of Race, Gender and Manner of Dress on Retail Service*, 11 MKTG. LETTERS 129, 133 (2000).

⁴⁸ Dean A. Dabney, Laura Dugan, Volkan Topalli & Richard C. Hollinger, *The Impact of Implicit Stereotyping on Offender Profiling: Unexpected Results from an Observational Study of Shoplifting*, 33 CRIM. JUST. & BEHAV. 646, 660–70 (2006).

⁴⁹ Schreer et al., *supra* note 18, at 1432–33.

⁵⁰ For a similar observation about the lack of available database from which to study retail race discrimination more generally, see GABBIDON & HIGGINS, *supra* note 18, xii, which notes that “scholars have generally neglected profiling that occurs in retail settings” as a “result of the fact that there is no easily accessible national database . . . from which to study the problem.” See also Zachary W. Brewster & Gerald Roman Nowak III, *Racialized Workplaces, Contemporary Racial Attitudes, and Stereotype Endorsement: A Recipe for Consumer Racial Profiling*, 64 SOCIO. PERSP. 343, 343 (2021) (observing that “investigations centered on everyday racial discrimination in frequently utilized consumer markets are surprisingly quite rare”).

⁵¹ Sophia R. Evett, Anne-Marie G. Hakstian, Jerome D. Williams & Geraldine R. Henderson, *What’s Race Got to Do with It? Responses to Consumer Discrimination*, 13 ANALYSES SOC. ISSUES & PUB. POL’Y 165, 168 (2013).

⁵² GABBIDON & HIGGINS, *supra* note 18, xii.

typically do not permit sales personnel to deviate from the standardized text of their agreements.⁵³ For instance, Harvard Law professor Todd Rakoff has observed that the “characteristics of firms counsel the adoption of standard forms and rigidify allegiance to them,” and that the use of standard form contracts keeps “wayward sales personnel” in check by ensuring adherence to the terms of the form contract.⁵⁴ Indeed, the widespread use of standard form contracts has led some experts to conclude that there is no longer any bargaining in consumer contracts.⁵⁵

However, more scholars are now suggesting that in competitive markets, sellers' enforcement of their contractual provisions may be more lenient than contractually required.⁵⁶ As these commentators have posited, sellers may include seemingly rigid, nonnegotiable terms in their contracts to protect themselves from consumer opportunism, while still allowing their representatives to deviate from standard terms on a case-by-case basis.⁵⁷

⁵³ See, e.g., Stewart Macaulay, *Private Legislation and the Duty to Read—Business Run by IBM Machine, the Law of Contracts and Credit Cards*, 19 VAND. L. REV. 1051, 1059 (1966) (observing that form contracts allow sellers to efficiently control salespeople by putting customers on notice of salespeople's “limited authority”); Samuel I. Becher & Esther Unger-Avram, *The Law of Standard Form Contracts: Misguided Intuitions and Suggestions for Reconstruction*, 8 DEPAUL BUS. & COM. L.J. 199, 201 (2010) (noting that the prevailing wisdom is that sellers are “not likely” to allow “deviations from pre-printed forms” and that “the typical seller does not empower its representatives (salespeople) to make changes in [a standard form contract]”).

⁵⁴ Todd D. Rakoff, *Contracts of Adhesion: An Essay in Reconstruction*, 96 HARV. L. REV. 1173, 1222–23 (1983); see also Todd D. Rakoff, *The Law and Sociology of Boilerplate*, 104 MICH. L. REV. 1235, 1236 (2006) (making a similar observation).

⁵⁵ See, e.g., Jason Scott Johnston, *The Return of Bargain: An Economic Theory of How Standard-Form Contracts Enable Cooperative Negotiation Between Businesses and Consumers*, 104 MICH. L. REV. 857, 858 (2006) (“Among attorneys, judges, and legal academics, there is virtual consensus that the widespread use by business firms of standard-form contracts in their dealings with consumers has completely eliminated bargaining in consumer contracts.”).

⁵⁶ See, e.g., Lucian A. Bebchuk & Richard A. Posner, *One-Sided Contracts in Competitive Consumer Markets*, 104 MICH. L. REV. 827, 830 (2006) (suggesting reputational costs may induce sellers to treat buyers more fairly than is contractually required); Avery Katz, *The Strategic Structure of Offer and Acceptance: Game Theory and the Law of Contract Formation*, 89 MICH. L. REV. 215, 281 (1990); Clayton P. Gillette, *Rolling Contracts as an Agency Problem*, 2004 WIS. L. REV. 679, 704–12 [hereinafter Gillette, *Rolling Contracts*] (suggesting that sellers may use a “contract clause that assigns an entitlement to the seller, but that the seller may underenforce where it is dealing with a good claimant”); Clayton P. Gillette, *Pre-Approved Contracts for Internet Commerce*, 42 HOUS. L. REV. 975, 977 (2005) [hereinafter Gillette, *Pre-Approved Contracts*] (making a similar observation); Johnston, *supra* note 55 (making a similar observation); Samuel I. Becher & Tal Z. Zarsky, *Minding the Gap*, 51 CONN. L. REV. 69, 99 (2019) (suggesting that there may be a gap between sellers' written agreements and their on-the-ground implementation).

⁵⁷ See, e.g., Gillette, *Pre-Approved Contracts*, *supra* note 56, at 977 (observing that sellers may use “ostensibly oppressive terms” to allow themselves “discretion to treat buyers who appear to be acting in good faith differently from those who appear to be acting opportunistically”). Some commentators have

Empirical evidence from non-retail arenas supports this theory. For example, survey research suggests that credit card issuers—who typically include a “late fee” provision in their contracts—sometimes authorize agents to waive those fees when borrowers demonstrate that they accidentally missed a payment.⁵⁸ Similarly, airlines, whose “contracts of carriage” require that passengers pay a fee if they miss their flight, often authorize agents to waive the fee for certain passengers who miss a flight due to unforeseen circumstances.⁵⁹ And mortgage service providers, who have the contractual right to foreclose on defaulting borrowers, frequently authorize their agents *not* to exercise the right to foreclose on a defaulting borrower, depending, for example, on a borrower’s credit risk.⁶⁰

Drawing on this evidence, some scholars have speculated that salespeople might selectively enforce contract terms against certain consumer groups that are viewed as “less valuable” or “less sophisticated.”⁶¹ For example, law professors Samuel Becher and Tal Zarsky have hypothesized that “uninformed and weak groups of consumers” will be disadvantaged, while “sophisticated and informed” groups will be treated more forgivingly or generously.⁶² Hebrew University Law professor Eyal Zamir has similarly theorized that reputational forces “are much more likely to work in favor of large, recurring, and sophisticated customers—whose goodwill the supplier values highly—than in favor of the weak, occasional, and unsophisticated customers, whose goodwill is valued less.”⁶³

An observational study by University of Michigan Law professor Manisha Padi supports these assumptions by providing empirical evidence suggesting that the practice of deviating from standard contractual terms

also referred to the specific context of retail returns in this regard. *See, e.g.*, Johnston, *supra* note 55, at 873–74 (“[R]etailers . . . generally granted their on-the-ground employees vast amounts of discretion in liberalizing their official return policies.”).

⁵⁸ *See, e.g.*, Michelle Crouch, *Poll: You Can Get Better Credit Card Terms Just by Asking*, YAHOO! FIN. (Mar. 27, 2017), <https://finance.yahoo.com/news/poll-better-credit-card-terms-030000309.html> [<https://perma.cc/4X48-FWZQ>]; Sophia Acevedo, *Avoiding Bank Fees: A Comprehensive Review*, BUS. INSIDER (July 23, 2024), <https://www.businessinsider.com/personal-finance/how-to-avoid-bank-fees> [<https://perma.cc/4QGY-BQKS>] (suggesting fees may be negotiated by consumers with agents).

⁵⁹ *See, e.g.*, Claire Nowak, *This Little-Known Airplane Rule Can Help when You Miss Your Flight*, READER’S DIG. (Mar. 6, 2019), <https://www.rd.com/advice/travel/missed-flight-flat-tire-rule/> [<https://perma.cc/4KNZ-KPHB>].

⁶⁰ *See* Robert B. Avery, Raphael W. Bostic, Paul S. Calem & Glenn B. Canner, *Credit Risk, Credit Scoring, and the Performance of Home Mortgages*, 82 FED. RES. BULL. 621, 622 (1996).

⁶¹ *See, e.g.*, Becher & Zarsky, *supra* note 56, at 91 (describing the “negative distributive problem” that arises when firms engage in ex post discrimination); Eyal Zamir, *Contract Law and Theory: Three Views of the Cathedral*, 81 U. CHI. L. REV. 2077, 2100 (2014).

⁶² Becher & Zarsky, *supra* note 56, at 91.

⁶³ Zamir, *supra* note 61, at 2100.

in mortgage agreements may disproportionately benefit those in better socioeconomic positions.⁶⁴ Using a detailed commercial dataset of mortgage performance from 2000 to 2008, Padi showed that 40% of the borrowers who fell behind on their payments during this period avoided foreclosure due to the service provider's exercise of discretion,⁶⁵ and that borrowers from high-income neighborhoods had a 3.3% higher chance of avoiding foreclosure compared to borrowers from low-income neighborhoods.⁶⁶ Padi's study thus suggests that discretionary deviations in consumer contracts might lead to disproportionately favorable outcomes for higher income consumers.⁶⁷

Still, the question remains: Should we expect to see differential implementation of sellers' form contracts along racial lines? The argument that we should is consistent with the wealth of research investigating how racial discrimination manifests in day-to-day discretionary decisions across various domains.

Evidence from diverse fields, including law enforcement, judicial proceedings, housing, and employment, consistently shows that racial bias infiltrates human judgments.⁶⁸ This body of research underlines a troubling correlation: the greater the discretionary power in decision-making, the larger the racial disparities observed.

For instance, in the realm of employment, a study simulating hiring processes revealed a pronounced bias against Black candidates when their qualifications were unclear, requiring recruiters to rely more heavily on discretion.⁶⁹ Similarly, in the legal system, Harvard Law professor Crystal Yang examined the consequences of the Supreme Court's *United States v. Booker* decision, which relaxed mandatory sentencing guidelines and increased judicial discretion in sentencing.⁷⁰ Using data spanning from

⁶⁴ Manisha Padi, *Contractual Inequality*, 120 MICH. L. REV. 825, 875 (2022).

⁶⁵ *Id.* at 830.

⁶⁶ *Id.* at 858.

⁶⁷ Padi conducted her analysis by neighborhood, rather than at the individual level, so she could not untangle the effects of income and race. But her research suggests that further study is warranted, since socioeconomic status and zip codes are often closely associated with race (e.g., zip codes with mean incomes lower than \$100,000 a year having a larger Black population than zip codes with higher mean incomes). Padi acknowledges this issue and discusses future research directions. *Id.* at 862.

⁶⁸ For a general overview of this vast literature, see Bertrand & Duflo, *supra* note 11.

⁶⁹ John F. Dovidio & Samuel L. Gaertner, *Aversive Racism and Selection Decisions: 1989 and 1999*, 11 PSYCH. SCI. 315, 318 (2000).

⁷⁰ Crystal S. Yang, *Free at Last? Judicial Discretion and Racial Disparities in Federal Sentencing*, 44 J. LEGAL STUD. 75 (2015).

1994 to 2010, her research demonstrates a marked increase in the sentencing disparity between Black and white defendants post-*Booker*.⁷¹

Indeed, the literature already widely acknowledges that “unstructured decision-making is exactly the sort of environment in which implicit biases can have their biggest impact.”⁷² As University of Oregon Law professor Erik Girvan explains, “implicit bias tends to influence decisions that are inherently ambiguous, difficult, or subjective,” as these situations typically require decision-makers to exercise their discretion, make a judgment call, or rely on intuition.⁷³

Extrapolating from these insights to the retail sector, it is plausible that store agents, when vested with the discretion to interpret and apply store policies or contractual terms, may do so in ways that systematically disadvantage Black customers. This Article introduces a pioneering audit study aimed at examining this theory within the context of retail return policies.

II. THE CASE STUDY AND METHODOLOGICAL APPROACH

A. *The Case Study: Receipt-Only Retail Returns*

This study examined racial disparities in the enforcement of retail return policies. Given the significant discretion store agents often have in accepting

⁷¹ *Id.* at 77. But see Joshua B. Fischman & Max M. Schanzenbach, *Racial Disparities Under the Federal Sentencing Guidelines: The Role of Judicial Discretion and Mandatory Minimums*, 9 J. EMPIRICAL LEGAL STUD. 729 (2012) (finding that increased judicial discretion following the shift to nonbinding sentencing guidelines did not necessarily result in greater racial disparities in sentencing post-*Booker*).

⁷² Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan & Cass R. Sunstein, *Discrimination in the Age of Algorithms*, 10 J. LEGAL ANALYSIS 113, 120 (2018).

⁷³ Erik J. Girvan, *When Our Reach Exceeds Our Grasp: Remedial Realism in Antidiscrimination Law*, 94 OR. L. REV. 359, 375 (2016). The substantiated claim that broader discretion might result in greater discrimination has also been raised and discussed in the context of the “rules versus standards” conundrum. See, e.g., Cass R. Sunstein, *Problems with Rules*, 83 CALIF. L. REV. 953, 974 (1995) (observing that rules can counteract bias, favoritism, or discrimination because they are “associated with impartiality”). For a general overview of the rules versus standards debate, see, for example, Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557, 559–60 (1992); Duncan Kennedy, *Form and Substance in Private Law Adjudication*, 89 HARV. L. REV. 1685, 1701 (1976); and Russell D. Covey, *Rules, Standards, Sentencing, and the Nature of Law*, 104 CALIF. L. REV. 447, 450 (2016).

or denying returns,⁷⁴ and the high value consumers place on this flexibility,⁷⁵ retail returns offered an ideal window into potential discriminatory practices in contractual performance.

The focus is on “receipt-required” return policies, ubiquitous across retail settings and often criticized by consumers as overly burdensome.⁷⁶ These policies primarily aim to safeguard retailers from fraud, such as shoplifting or the return of items purchased elsewhere, presenting a dual challenge: deterring fraudulent returns while avoiding alienation of genuine customers.⁷⁷ Therefore, the test case of retail returns presents an ideal setting for studying discretionary practices and ascertaining if they reveal racial disparities in the treatment of customers.

B. The Methodological Approach

The study's primary goal was to determine whether salespeople are more inclined to bend return policies for white customers compared to Black

⁷⁴ See, e.g., Bebachuk & Posner, *supra* note 56, at 833–34 (noting sellers often accept returns in circumstances “in which they are not obligated to do so”); Johnston, *supra* note 55, at 873–74 (“[R]etailers . . . generally granted their on-the-ground employees vast amounts of discretion in liberalizing their official return policies so as to please consumers.”); Gillette, *Rolling Contracts*, *supra* note 56, at 706; Becher & Zarsky, *supra* note 56, at 91. Additionally, in informal interviews conducted with store clerks in preparation for this study, several interviewees also mentioned being granted a considerable degree of discretion in deciding whether to accept a return. These interviews are on file with *Northwestern University Law Review*.

⁷⁵ According to recent data, over 15% of purchases in the United States are returned annually. See Melissa Repko, *A More than \$761 Billion Dilemma: Retailers' Returns Jump as Online Sales Grow*, CNBC (Jan. 25, 2022, 2:02 PM), <https://www.cnbc.com/2022/01/25/retailers-average-return-rate-jumps-to-16point6percent-as-online-sales-grow-.html> [<https://perma.cc/5R55-ZQXM>]; see also EYAL ZAMIR & DORON TEICHMAN, *BEHAVIORAL LAW AND ECONOMICS* 290–91 (2018) (observing that many consumers value the right to return retail products to stores because they often regret the purchase after the fact).

⁷⁶ In a larger sample of 192 retail stores (established for a separate study), receipt requirements were mentioned in 84% of the return policies. This data is on file with the author. For evidence that receipt requirements are perceived as a hassle by consumers, see, for example, Narayan Janakiraman, Holly A. Syrdal & Ryan Freling, *The Effect of Return Policy Leniency on Consumer Purchase and Return Decisions: A Meta-Analytic Review*, 92 J. RETAILING 226, 228 (2016), which notes that no-receipt return policies require less effort on the part of the consumer and are thus considered “more lenient.”

⁷⁷ See, e.g., Narayan Janakiraman, Holly A. Syrdal & Ryan Freling, *How to Design a Return Policy*, HARV. BUS. REV. (Sept. 9, 2016), <http://hbr.org/2016/08/how-to-design-a-return-policy> [<http://perma.cc/XT33-8HKF>] (discussing the trade-off between encouraging customers to shop at the store and discouraging customers from making returns); Danni Zhang, Regina Frei, P.K. Senyo, Steffen Bayer, Enrico Gerding, Gary Willis & Adrian Beck, *Understanding Fraudulent Returns and Mitigation Strategies in Multichannel Retailing*, J. RETAILING & CONSUMER SERVS., Jan. 2023, at 1, 10–11 (“One of the most effective technologies to prevent returns fraud is the use of unique product identifiers (UIDs), which are then linked to the purchase and shown on the receipt.”).

customers requesting similar concessions.⁷⁸ The study tested the hypothesis that white customers are more likely to receive lenient treatment than their Black counterparts in comparable situations. It also explored how race and gender intersect in the application of retail return policies.

To investigate these issues, I first employed an audit methodology, sending testers with matching observable characteristics, except for race and gender, to various retail outlets. Audit studies, commonly referred to as “field experiments,” are highly regarded in social science for establishing causal inferences in real-world settings.⁷⁹ Essentially, they allow researchers to conduct a controlled experiment in real-life environments. By carefully controlling for the testers’ observable characteristics, audit studies ensure that any differences in treatment (such as access to goods or services) can be attributed to the variable of interest—in this case, race or gender. This level of control helps eliminate the impact of other factors that might otherwise confound the results, which is a common challenge in traditional regression analyses. In regression analyses, which rely on observational data rather than random assignment, there is a risk of “omitted variable bias”—where the effect of an unmeasured variable is mistakenly attributed to the measured variables. Field experiments avoid this problem by keeping all variables constant except the one being tested, ensuring a more accurate measurement of the causal effect.⁸⁰

⁷⁸ The study was approved by the University of Chicago Institutional Review Board and preregistered in AsPredicted (“Post-Contract Discrimination in the Retail Market,” #16928, created Nov. 23, 2018).

⁷⁹ 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, 993–94 (2004) (discussing the advantages of field experiments over other statistical methods to study discrimination); cf. James J. Heckman, *Detecting Discrimination*, 12 J. ECON. PERSP. 101, 102, 107–11 (1998) (discussing the limitations of audit studies).

⁸⁰ See, e.g., Bertrand & Duflo, *supra* note 11, at 157 (noting that “audit . . . methodologies were developed to address [the] core limitations of the regression approach to measuring discrimination” and surveying the benefits of audit studies). Admittedly, audit studies are not without limitations. Ensuring controlled conditions in field settings demands meticulous planning and coordination. Ethical considerations also come into play, particularly regarding the impact on unsuspecting participants in the real world and the challenges around informed consent. However, while other research methods provide valuable insights—observational studies shed light on patterns and correlations, and lab experiments clarify causal mechanisms under controlled conditions—only audit or correspondence studies offer a blend of observing real behaviors in natural environments with the benefits of robust causal inference. For a detailed analysis of the advantages and limitations of audit and correspondence studies, see Netta Barak-Corren & Tamir Berkman, *Constitutional Consequences*, 99 N.Y.U. L. REV. 785, 804–05 (2024), which notes that “[f]ield experiments provide the most powerful method for causal inference in real life settings currently known to the social and legal sciences.”

Professors Ian Ayres and Peter Siegelman provide an illustrative example of the power of audit studies.⁸¹ As part of their 1995 study, they randomly assigned Black and white testers to negotiate car prices in various dealerships in Chicago.⁸² The findings showed a consistent pattern of white male shoppers receiving lower price quotes than their white female and Black counterparts. This discriminatory practice might have been less detectable through regression analysis, where identifying causal relationships typically poses a challenge.⁸³

In more contemporary research, audit studies on consumer market discrimination, often called correspondence studies, have utilized hypothetical customer inquiries via letters or emails to assess sellers' reactions, employing proxies such as names to infer customers' race, gender, or other characteristics.⁸⁴ Although correspondence studies can efficiently generate expansive datasets at lower costs,⁸⁵ they might not accurately reflect the extent of discrimination, particularly when they employ coarse proxies such as names to signal race.⁸⁶ Moreover, both methods generally only

⁸¹ Ayres & Siegelman, *supra* note 13.

⁸² *Id.* at 305.

⁸³ A key limitation of regression analysis is “omitted variable bias,” which arises when an important explanatory factor is excluded from the model. If the analysis fails to account for certain variables influencing pricing decisions—such as perceived income level, bargaining style, or product type—observed disparities may be attributed to race or gender when, in fact, other unmeasured factors could be driving the differences. This challenge is particularly pronounced in discrimination research, where regression estimates of racial or gender disparities can be difficult to interpret. *See id.* at 309–10; Bertrand & Duflo, *supra* note 11, at 157 (reviewing the limitations of regression analysis in studying race and gender discrimination, including the omitted variable bias, and observing that “[a]udit and correspondence methodologies were developed to address these core limitations of the regression approach to measuring discrimination”).

⁸⁴ For notable contemporary illustrations, see, for example, Edelman et al., *supra* note 42, at 7, which found, based on a correspondence study, that Airbnb hosts were more likely to refuse booking requests made by guests with Black-sounding names compared to guests with white-sounding names); Netta Barak-Corren, *A License to Discriminate? The Market Response to Masterpiece Cakeshop*, 56 HARV. C.R.-C.L. L. REV. 315, 341, 344–46 (2021); and Netta Barak-Corren, *Religious Exemptions Increase Discrimination Toward Same-Sex Couples: Evidence from Masterpiece Cakeshop*, 50 J. LEGAL STUD. 75, 95 (2021), which both found, based on a correspondence study, that bakers were more likely to refuse requests made by gay couples, as compared to heterosexual couples, using gendered names to signal same-sex couples.

⁸⁵ *See, e.g.,* Bertrand & Duflo, *supra* note 11, at 161; S. Michael Gaddis, *How Black Are Lakisha and Jamal? Racial Perceptions from Names Used in Correspondence Audit Studies*, 4 SOCIO. SCI. 469, 469 (2017) (“[C]orrespondence audits enable researchers to . . . collect larger samples and reduce research costs.”).

⁸⁶ Gaddis, *supra* note 85, at 469 (observing that in correspondence studies, as opposed to in-person audits, researchers “lose the ability to more directly convey race through appearance and interaction and instead rely solely on names to signal race”).

capture the early stages of transactions, examining outcomes such as response rates or the likelihood of a “call-back.”⁸⁷

This research contributes to existing literature on market discrimination in three significant respects. First, it studies seller behavior in the post-contractual phase by employing actual individuals to navigate the nuanced reality of racial identity—a complexity that the mere “racial soundingness” of a name cannot fully encapsulate.⁸⁸

A second methodological strength of the research methodology employed in this Article is the integration of interviews with the audit study. While audit studies are adept at detecting discrimination, they are less capable of uncovering the underlying motivations.⁸⁹ To address this limitation, I conducted interviews with decision-makers—store agents not involved in the audit study. While I expected that interviewees would not explicitly acknowledge prejudiced attitudes or discriminatory behaviors, with thoughtfully designed questions, I hoped to glean insights that align with the patterns observed in the audits. These insights might include the decision-makers’ interpretations based on a customer’s race and their perspectives on how race influences their decision-making processes. Such information is vital in understanding the context and motivations behind the behavior identified in the audit study. Qualitative data also enables us to understand structural pressures exerted on decision-makers, which then makes discriminatory practices more or less likely—insights invaluable in crafting effective policy responses.

This study makes a third significant methodological contribution by addressing the challenges inherent in analyzing *ex post* discrimination in contractual relationships. Typically, the inherent variability in contractual dynamics—from the preliminary stages of pre-contract negotiations to the execution of the contract—complicates the study of post-contract outcomes. Such variability obscures the link between consumers’ characteristics and their experiences after entering a contract. Consider mortgage foreclosures as an illustrative example: the decision of a mortgage agent to initiate foreclosure on a borrower in default could be swayed by their history of interactions, including those that occurred before contract formation. This

⁸⁷ Bertrand & Duflo, *supra* note 11, at 174 (“Yet another limitation of field studies . . . is that . . . [m]ost of the time, interview invitations or rental offers (‘call-back rates’) are the only outcomes captured by field experiments . . .”).

⁸⁸ See, e.g., Sen & Wasow, *supra* note 10, at 500 (noting that the technique of varying the “racial soundingness” of an applicant’s name introduces various challenges, as “racial soundingness” is not synonymous to race).

⁸⁹ See Bertrand & Duflo, *supra* note 11, at 176.

intertwining of past interactions makes it particularly difficult to isolate and understand the impact of specific consumer characteristics, such as race, on decisions such as foreclosure, independent of previous interactions. Thus, by sending testers to return items that had been purchased by different members of the research team, this study allows for the isolation of the impact of consumer characteristics on post-contract treatment, devoid of the noise introduced by earlier interactions. By doing so, it provides a more refined and accurate lens through which we can understand the dynamics of post-contract discrimination, deepening our comprehension of how the intrinsic traits of the parties in a contractual relationship might shape their experiences and outcomes, independent of their precontractual conduct.

III. THE AUDIT STUDY

A. *The Sample*

The initial sample of stores comes from the Data Axle Reference Solutions' (Reference Solutions') and Hoover's Company Directories' databases according to the following selection criteria: all had a downtown Chicago location, and all expressly required a receipt for refunds.⁹⁰ To keep the item prices fixed across stores at between \$20 to \$30, dollar stores or high-end designer stores that did not offer items for \$30 or less were excluded from the sample. Table 1 presents the sampled stores' summary statistics.

⁹⁰ Data Axle's Reference Solutions is a comprehensive business and consumer database providing detailed information on U.S. companies. Its primary use is for market research and competitive analysis. The database includes various segments such as U.S. businesses and the retail sector. Reference Solutions sources its data from a combination of publicly available business records, including phone directories, annual reports, SEC filings, and government data such as business registrations. Reference Solutions also verifies the accuracy of its listings through direct telephone verification and updates on a regular basis to maintain current information. *See* DATA AXLE REFERENCE SOLS., TECHNICAL OVERVIEW 3–24, <https://referencesolutions.data-axle.com/wp-content/uploads/2023/07/all-business-databases.pdf> [<https://perma.cc/4XQ4-YC63>]. Hoover's is a business information and research database providing detailed profiles on companies and industries worldwide. It offers users access to in-depth company data, including financials, industry trends, executive information, and market research. It is frequently used for sales prospecting, competitive analysis, and market research. Hoover's sources its data primarily from Dun & Bradstreet's extensive business database, which gathers information from a wide variety of sources. These include public financial reports, SEC filings, annual reports, press releases, and news articles. Hoover's also obtains data through direct outreach to companies, customer contributions, and proprietary analytics, ensuring it stays current. *See Put the World's Largest Database to Work for Your Business*, DUN & BRADSTREET, <http://www.dnb.com/en-us/products/try/dnb-hoovers-free-trial-business-database.html> [<https://perma.cc/WS86-EF93>].

TABLE 1: SAMPLED STORES' DESCRIPTIVE STATISTICS

	Mean	Median	Standard Deviation	Min	Max
Company's Annual Revenues	176.98	3.15	870.48	0.0028	5575.44
Company Age (years)	62.8	46.5	43.87	5	201
Company Is Publicly Traded	0.66	1	0.47	0	1
Median Price of Store Items (in USD)	60	45	59.7	5	350
Company Has Lenient Formal Policy	0.635	1	0.48	0	1

Note. Annual revenues are in millions of U.S. dollars. Age is defined as the years passed since the company's incorporation. The median price of the store items refers to the median price of the items offered for sale at the store's website. "Lenient formal policy" is defined as a policy that formally allows store clerks to offer consumers non-refund concessions (store credit or exchange) even when they fail to show a receipt.

B. The Design of the Audits

The audit-study team included three "buyers" and nineteen "testers," all of whom were undergraduate students recruited from local universities and trained by the research team. At the first stage, the "buyers" were instructed to purchase gender-neutral accessories (i.e., hats, scarves, gloves, or socks)⁹¹ from fifty-nine selected retail clothing stores in downtown Chicago.⁹² At the second stage, beginning a week after the completion of the first stage, the testers then attempted to return the items and receive a refund without a receipt, despite the fact that the store's official policy required a receipt as a condition to a refund.⁹³ All items were in their original packaging and

⁹¹ If no accessories were available, the buyers were instructed to purchase a shirt, pair of pants, or another clothing item. Admittedly, some of the accessories might have been gendered despite these instructions. Future studies should attempt to more closely control for this issue.

⁹² The buyers were specifically instructed to refrain from purchasing personal items (e.g., underwear, swimwear, or jewelry), electronic devices, clearance or sale items, or any item that was specifically not eligible for returns according to each store's formal return policy. To eliminate concerns that the cost of the returned item would drive any observed disparities, purchased items' prices were kept constant (i.e., between \$20 and \$30). Moreover, the buyers paid in cash so that sellers would not obtain any personal information from the purchase.

⁹³ This study's design, in which the testers returned items that had been purchased in advance by the buyers, was chosen for two main reasons. First, it allowed for the purchase of almost identical items from each store. Second, it mitigated the concern that store agents might identify the person making the return. The time that elapsed between the purchase and the attempted return also alleviated the concern that store

condition and with tags attached. The testers—five white women, five white men, four Black women,⁹⁴ and five Black men—were randomly assigned to stores, and the order in which testers audited each store was also chosen randomly.

To reinforce confidence that unobserved differences between testers would not drive any differences in outcomes, recruited testers met uniform criteria: all were undergraduate students residing in Chicago, all were 18–25 years of age, and all wore similar attire (casual clothing).⁹⁵ In addition, all testers were required to memorize (and then practice) an identical script and then follow it exactly in their interactions with store agents.⁹⁶ Testers were provided with uniform answers about the reason for the return and for not having the receipt.⁹⁷ They also received a list of contingent responses to additional questions they were likely to encounter.⁹⁸ To avoid any conscious or unconscious bias on the part of the tester, testers were not told the true purpose of the study. Rather, they were told that the study's purpose was to explore differences in retailers' return policies.⁹⁹

At each store, the testers were instructed to wait in line until a store clerk became available, inform the clerk that they wanted to return the item, and await the store clerk's response. If the store clerk *immediately* agreed to provide a refund, testers were instructed to accept the refund, thank the clerk,

agents would remember that the person making the return was not the person making the original purchase.

⁹⁴ I initially hired five Black women as testers. However, one tester dropped out during training.

⁹⁵ The study endeavored to mitigate class and socioeconomic biases by standardizing attire among all testers. Yet, it is crucial to recognize that identical clothing can project disparate connotations contingent upon the racial or ethnic identity of the wearer. For a comprehensive exploration of this phenomenon and its implications, see *infra* text accompanying notes 130–132. Also, before sending the testers to the stores, the testers verified they had not attended the stores on their list in the preceding three months to avoid concerns that the store agents would recognize or be familiar with them.

⁹⁶ Before auditing the stores, the testers attended training sessions at the University of Chicago, where they practiced their scripts and participated in numerous mock negotiations designed to help them behave uniformly during the audits.

⁹⁷ For example, if the store agent asked why they wanted to return the item, testers were instructed to say that they realized they did not need it after purchasing it.

⁹⁸ It is possible that store agents' responses to testers' return requests were sensitive to the reason offered by testers. For example, if testers said that they bought the wrong size or received the item as a gift and did not like it, store agents may have responded differently. The generic excuse used in the study was intended to allow testers to request a refund rather than merely exchange the item or obtain store credit. Unlike returning a gift or exchanging an item for a different size, explicitly saying that they did not need the product made asking for a cash refund rather than an exchange or store credit arguably more credible and reasonable. In any event, the type of excuse chosen cannot, on its own, explain the observed racial disparities in treatment.

⁹⁹ In addition, testers were asked at the end of audits what they thought the study was about, and none of them indicated they believed the study tested for discrimination or racial bias.

and leave the store. If, however, testers were told they could not return the item or were offered anything other than a refund (for example, an exchange or store credit), they were instructed to ask to speak to a manager. Whether the store clerk refused to call someone or whether the store clerk called a manager, testers were instructed to once again ask for a refund. Regardless of the clerk's or manager's response, testers would then offer thanks and leave the store.

Upon leaving the store, the testers were instructed to fill out a detailed report of their audit, including: the date and time of the visit, the perceived race and gender of the store clerk (and manager if needed), and—most importantly—the store clerks' (or managers') immediate and final responses to their return requests (which they were asked to classify as “refund,” “store credit,” “exchange,” “complete refusal to accept the return,” or “other” with an explanation).¹⁰⁰

In designing the study, I recognized that immediate outcomes would probably represent a “cleaner” test of discrimination, as they required considerably less interaction between the testers and the store agents.¹⁰¹ By including the post-complaint results in the analysis, however, I was able to test whether complaining or negotiating moderated the observed racial disparities. After all, one could hypothesize that, even if we observed racial disparities in treatment after the initial return request, these disparities should decrease (or even disappear) once the tester lodges a complaint and again requests a refund. In other words, examining eventual outcomes acted as a robustness check on the study's findings.

¹⁰⁰ The results were then coded and analyzed using both points in time during the tester's interchange(s) at the store: first, the store clerk's immediate response to the request for a refund (“immediate outcomes”), and second, the store clerk's or manager's final response to the request for a refund (“eventual outcomes”). In cases where the testers were immediately offered a cash refund, both their immediate and eventual return outcomes were coded as a “refund.”

¹⁰¹ Because it is still possible that inter-tester variation other than race or gender produced some of the disparities, even at the immediate stage, leave-one-out cross-validations were conducted to ensure that the results of any particular tester did not drive the observed disparities. The cross-validations confirmed that any one tester did not drive the results. Jackknife and bootstrap resampling operations were also conducted to ensure robustness. Resampling is a method of reusing data to generate new, hypothetical samples (called resamples) that are representative of the underlying population. While both bootstrap and jackknife could be used as resampling methods, jackknife is a more conservative approach and is considered more suitable for small original data samples. The jackknife is used to estimate the bias of an estimator by dropping one observation each time and recalculating the estimator based on the remaining dataset. The results generally remain statistically significant when the standard errors are calculated using the jackknife and bootstrap operations. These analyses are on file with the author.

C. Results

1. Key Findings: Racial Bias in Immediate Return Outcomes

This Section presents the main findings, using the tester's "immediate return outcome"—defined as the store clerk's first response to the tester's return request—as the main outcome measure. Because store clerks' immediate reactions occurred after minimal interaction with the testers, this measure offers a well-controlled test for discriminatory behavior. In other words, it is unlikely that racial disparities in immediate treatment resulted from uncontrolled differences in tester behavior.

Figures 1A and 1B below visualize the raw results. Figure 1A breaks the testers' immediate return outcomes into three categories: "offers of refunds," "store credits or exchanges" (excluding refunds), and outright denials of return requests ("rejections"). Figure 1B presents the proportion of immediate refund offers (left panel) and immediate return rejections (right panel) based on the tester's race and gender (White Female, White Male, Black Female, Black Male). Error bars represent 95% confidence intervals.

FIGURE 1A: IMMEDIATE RETURN OUTCOMES BY TESTER RACE AND GENDER

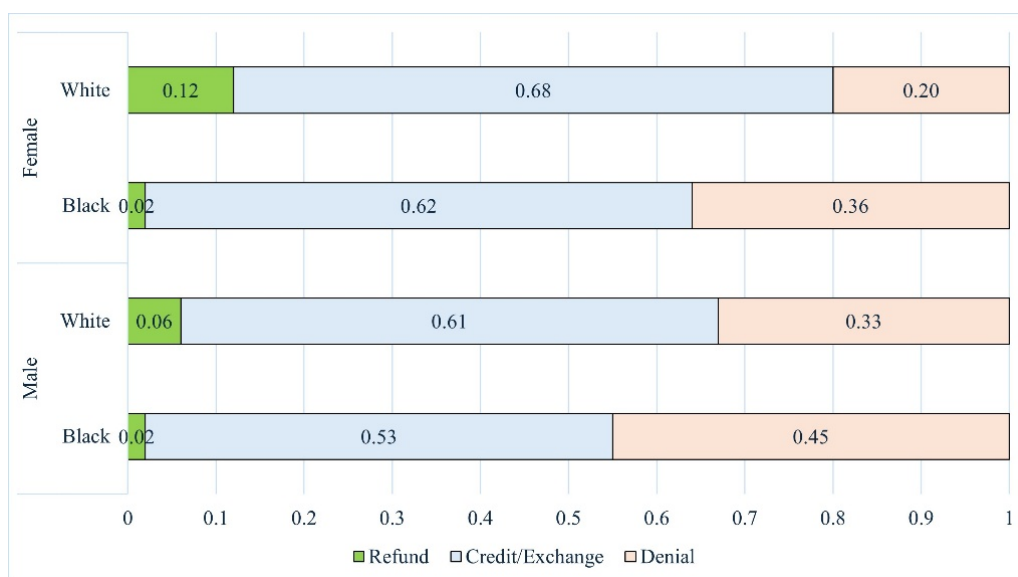
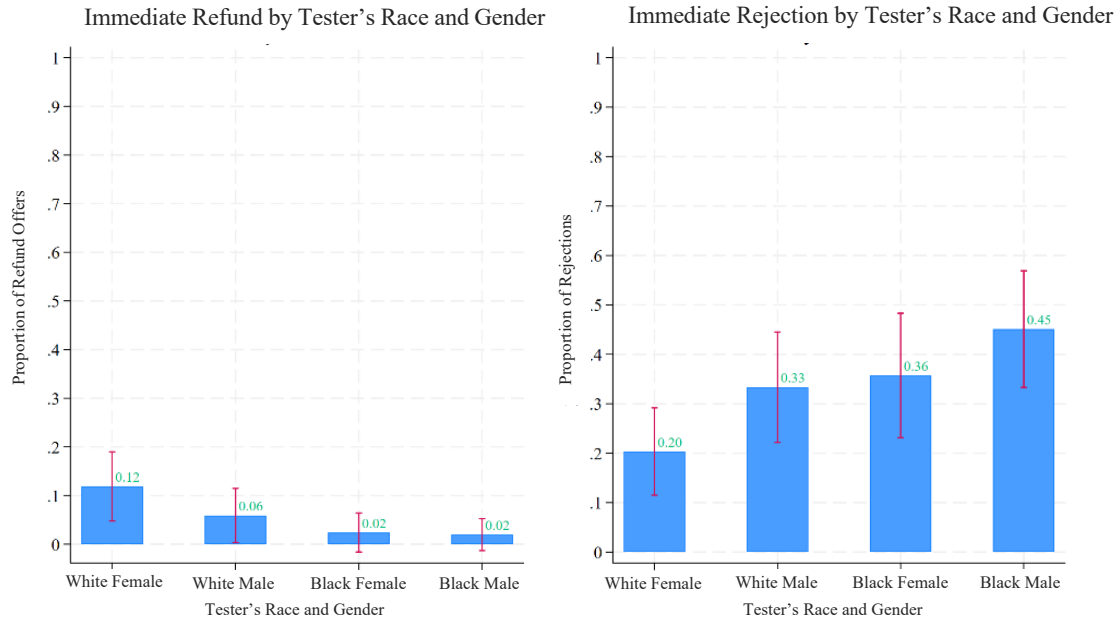


FIGURE 1B: IMMEDIATE RETURN OUTCOMES BY TESTER RACE AND GENDER



As Figures 1A and 1B show, white testers received refund offers at higher rates than Black testers within each gender subgroup (12% for white females versus 2% for Black females; 6% for white males versus 2% for Black males). They also faced lower rejection rates (20% for white females versus 36% for Black females; 33% for white males versus 45% for Black males). However, these raw results do not account for other factors that may have influenced tester treatment, such as the time and day of the audit, as well as clerk and store characteristics. To address these variables, I conducted regression analyses.

In the regression analysis that follows, I consider two dependent variables: (1) “Immediate refund,” a binary variable coded as “1” if the tester received a refund offer immediately after the initial return request, and “0” otherwise; and (2) “Immediate rejection,” a binary variable coded as “1” if the return request was denied outright, and “0” if any immediate concession (e.g., refund, store credit, or exchange) was granted.

Since almost all stores were audited multiple times,¹⁰² and some shared a parent company,¹⁰³ all regression models include both store and parent company fixed effects.¹⁰⁴ Additionally, day-of-audit and time-of-audit fixed effects are incorporated to control for potential variations in treatment based on audit timing.

Regression Table 1 presents the results.¹⁰⁵ As the Table illustrates, the findings revealed significant differences in the immediate treatment of Black and white shoppers. Controlling for gender, white shoppers were about 7.8 percentage points more likely to receive an immediate refund and about 14.5 percentage points less likely to have their return request outright rejected (without even an offer of store credit or exchange) than their Black counterparts.¹⁰⁶ Figure 2 below visualizes these findings.

Observed racial differences in treatment remain statistically significant and become even larger in magnitude once the perceived race and gender of the store clerk is controlled for in the regression, with white testers (controlling for gender) about 8.9 percentage points more likely to receive a refund offer and about 16.3 percentage points less likely to have their request rejected completely compared to Black testers.¹⁰⁷

¹⁰² Only two of the sampled stores were audited once, while the remaining fifty-seven stores were audited several times: seven stores underwent two audits, thirteen stores experienced three audits, and the majority—thirty-seven stores, constituting approximately 65% of the sample—were audited four times, each by a different tester (a white male, a white female, a Black male, and a Black female).

¹⁰³ In the dataset, twenty-five of the audited stores are clustered within eleven parent companies, with each parent company encompassing between two to three stores. Conversely, the remaining thirty-two stores are individually associated with distinct parent companies, illustrating a diverse and varied corporate structure within the audited entities.

¹⁰⁴ Because some observations in the data come from the same store, there might be unobserved characteristics that are specific to each store, but that affect all the observations from that store. Adding store fixed effects accounts for the unobserved characteristics that are constant within each store over time, helping to isolate the effect of the study's variables of interest and control for any potential bias caused by store-specific factors. In addition to store fixed effects, parent company fixed effects account for differences between companies that own multiple stores. Even if two stores are different, they might share certain corporate-level policies or business practices. Including parent company fixed effects helps control for the influence of these broader company-level factors.

¹⁰⁵ To validate these findings, I also conducted analyses using logistic regression models. These analyses corroborated the results. *See infra* Online Appendix, Table 1A. As shown in columns (1) and (5), white testers were significantly more likely to receive a refund offer and significantly less likely to face complete rejection of their return request than comparable Black testers. These differences are statistically significant (at the 10% and 5% levels) even when using logistic regressions instead of simple OLS regressions.

¹⁰⁶ See columns (1) and (5), respectively. Both differences were statistically significant at the 5% level ($n = 197$ observations).

¹⁰⁷ See columns (3) and (7), respectively. Both differences were statistically significant. Differences in the likelihood of receiving refund offers were significant at the 10% level, and differences in the likelihood of facing immediate rejections were significant at the 5% level ($n = 197$ observations).

Notably, no significant interactions were found between the testers' race and gender and the perceived race and gender of the store clerks.¹⁰⁸ This suggests that the observed disparities in treatment are not driven by any particular race–gender subgroup of clerks.

REGRESSION TABLE 1: IMMEDIATE OUTCOMES BY TESTER RACE & GENDER
AND OTHER VARIABLES

Immediate Refund as Dependent Variable				
	(1)	(2)	(3) + Agent Demographics	(4) + Interactions
White tester	0.0782** (0.0324)	0.0660 (0.0440)	0.0889* (0.0326)	0.0966 (0.0721)
Female tester	0.0700* (0.0365)	0.0566 (0.0491)	0.0574 (0.0363)	0.0390 (0.0818)
White tester # Female tester		0.0285 (0.0691)		0.00186 (0.105)
White clerk			0.0703* (0.0398)	0.0593 (0.0851)
Male clerk			0.0624 (0.0407)	0.136 (0.107)
BF tester # BM clerk				0.0184 (0.155)
BM tester # WM clerk				−0.198 (0.153)
BF tester # WF clerk				0.0267 (0.124)
BF tester # WM clerk				−0.103 (0.186)
WM tester # BM clerk				−0.178 (0.131)
WF tester # BM clerk				−0.0806 (0.136)
WM tester # WF clerk				−0.0371 (0.126)

¹⁰⁸ See columns (4) and (8).

WM tester # WM clerk	0.0239 (0.214)			
WF tester # WF clerk	-0.0362 (0.119)			
WF tester # WM clerk	0.0462 (0.183)			
Constant	-0.0207 (0.0291)	-0.0156 (0.0317)	-0.0686* (0.0378)	-0.0591 (0.0574)
Observations	197	197	197	197
R²	0.481	0.482	0.506	0.539

Note. Immediate refund offers are used in models 1 to 4 as the dependent variable. Model 1 regresses the dependent variable on the race and gender of the tester. Model 2 adds an interaction term between the race of the tester and the gender of the tester. Model 3 adds the race and gender of the clerk (as perceived and reported by the tester) as controls. Model 4 adds a four-way interaction between the demographics (race and gender) of the tester and those of the clerk (as perceived and reported by the tester). All models include store, parent company, day-of-audit, and time-of-audit fixed effects. Standard errors are presented in parentheses. * $p < 0.1$ (significant at the 10% level), ** $p < 0.05$ (significant at the 5% level), *** $p < 0.01$ (significant at the 1% level).

NORTHWESTERN UNIVERSITY LAW REVIEW

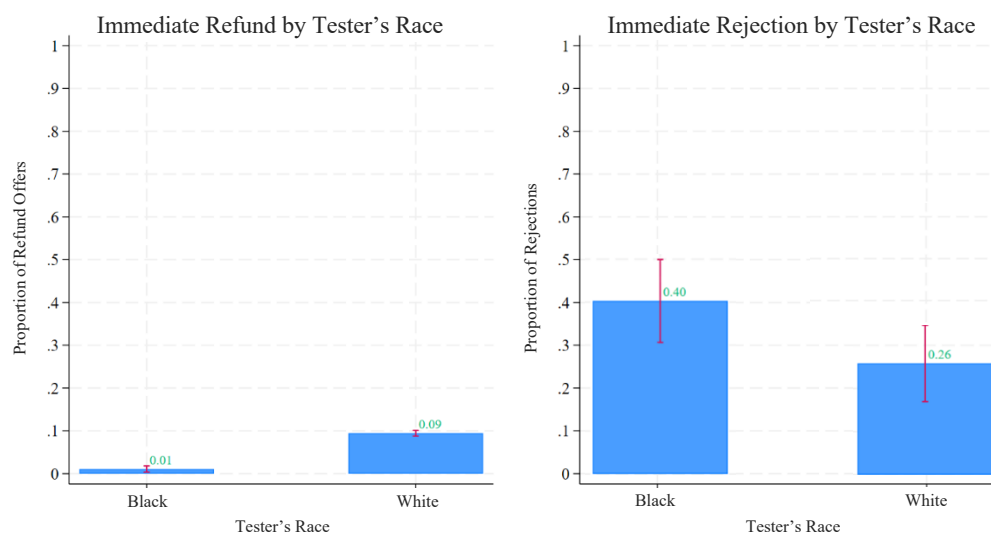
REGRESSION TABLE 1 (CONT.): IMMEDIATE OUTCOMES BY TESTER RACE & GENDER
AND OTHER VARIABLES

Immediate Rejection as Dependent Variable				
	(5)	(6)	(7)	(8)
White tester	-0.145** (0.0705)	-0.193** (0.0955)	-0.163** (0.0714)	-0.313** (0.157)
Female tester	-0.123 (0.0794)	-0.176 (0.107)	-0.099 (0.079)	-0.108 (0.178)
White tester # Female tester		0.113 (0.150)		0.0212 (0.229)
White clerk			-0.123 (0.0872)	-0.206 (0.186)
Male clerk			-0.12 (0.0892)	-0.321 (0.233)
BF tester # BM clerk				-0.319 (0.338)
BM tester # WM clerk				0.272 (0.333)
BF tester # WF clerk				0.0776 (0.270)
BF tester # WM clerk				0.0334 (0.406)
WM tester # BM clerk				0.292 (0.285)
WF tester # BM clerk				0.481 (0.297)
WM tester # WF clerk				0.0738 (0.275)
WM tester # WM clerk				0.397 (0.468)
WF tester # WF clerk				0.0976 (0.260)
WF tester # WM clerk				0.334 (0.400)

Constant	0.463*** (0.0632)	0.483*** (0.0688)	0.587*** (0.0824)	0.646*** (0.125)
Observations	197	197	197	197
R²	0.410	0.412	0.435	0.472

Note. Models 5 to 8 present the same four linear regressions as models 1 to 4, only this time using immediate rejections (i.e., refusals to offer refund, store credit, or an exchange) as the dependent variable. All models include store, parent company, day-of-audit, and time-of-audit fixed effects. Standard errors are presented in parentheses. * $p < 0.1$ (significant at the 10% level), ** $p < 0.05$ (significant at the 5% level), *** $p < 0.01$ (significant at the 1% level).

FIGURE 2: IMMEDIATE OUTCOMES BY TESTER'S RACE



Note. Figure 2 presents the predicted probabilities of immediate refund offers and rejections following a return request, disaggregated by tester race. The estimates are based on a regression model that controls for tester's gender and includes fixed effects for store, parent company, day of audit, and time of audit (columns (1) and (5) in Regression Table 1). The bars represent the adjusted mean probabilities of refund offers and rejections for Black and white testers, with error bars indicating 95% confidence intervals.

The results revealed that white female testers had a higher likelihood of securing refund offers compared to Black testers, both men and women. White women's chances of receiving a refund were approximately 10 percentage points greater than those of Black women and 14 percentage

points greater than those of Black men.¹⁰⁹ Regarding outright rejections, white women faced about 8 percentage points fewer denials compared to Black women and roughly 25 percentage points fewer than Black men.¹¹⁰ Additionally, white male testers outperformed Black male testers in terms of fewer outright denials, with a difference of 19 percentage points.¹¹¹ Notably, however, the return outcomes for white male testers were not significantly different from those for Black female testers.

2. *Is the “Race Effect” Influenced by Store Characteristics?*

To shed further light on this study’s findings, I investigated whether the observed preferential treatment towards white customers was influenced by various store characteristics. Specifically, I examined the interaction between the race of the tester and the following factors: the pricing levels at the store,¹¹² whether the store was part of a public or private corporation, and the annual revenues of the company. Regression Table 2 below shows these results.¹¹³ The analysis revealed no significant interaction between the race of the tester and any of these store or company attributes. This suggests that the racial disparities identified in this study are consistent and pervasive and are largely unaffected by the nature of the store or the financial metrics of the company involved.

¹⁰⁹ These differences are statistically significant at the 10% and 1% levels, respectively. White female testers were also more likely than white male testers to receive a refund offer, but this difference (9 percentage points) was not statistically significant.

¹¹⁰ Note, however, that while the difference between white female testers and Black male testers was significant at the 5% level, the difference between white female testers and Black female testers did not reach statistical significance.

¹¹¹ The difference between white male and Black male testers in terms of the likelihood of experiencing an outright denial was significant at the 10% level. White male testers were also 4 percentage points more likely to receive a refund offer compared to Black male testers, but this difference was not statistically significant ($p < 0.11$).

¹¹² Python was utilized to scrape median prices from store websites, opting for median over mean to reduce skew from outliers. Where online data was inaccessible, I manually gathered prices or selected 25 in-store items and computed their median prices.

¹¹³ Regression Table 2 uses immediate refund as the dependent variable. There are also no significant interactions between store or company characteristics and the race of the tester when immediate rejection is used as the dependent variable. *See infra* Online Appendix, Table 2A.

REGRESSION TABLE 2: IS THE RACE EFFECT INFLUENCED BY STORE CHARACTERISTICS?

	(1) + Store Median Prices	(2) + Store Is Privately Owned	(3) + Revenues
Black female tester	0.0635 (0.0479)	0.0594 (0.0480)	0.0476 (0.0466)
White male tester	0.0365 (0.0549)	0.0511 (0.0602)	0.0622 (0.0432)
White female tester	0.0841 (0.0580)	0.127** (0.0634)	0.144*** (0.0472)
White agent	0.0298 (0.0391)	0.0628 (0.0380)	0.0488 (0.0359)
Male agent	0.0585 (0.0412)	0.0553 (0.0380)	0.0521 (0.0388)
Median price	0.000493 (0.00970)		
Median price # White tester	0.000575 (0.000591)		
Store is privately owned		-0.0921 (0.0645)	
Public store # White tester		0.0474 (0.0627)	
Revenues			0.0109** (0.00440)
Revenues # White tester			0.00000135 (0.0000322)
Constant	-0.00439 (0.176)	0.0729 (0.186)	-0.106 (0.163)
Observations	186	203	185
R²	0.507	0.488	0.536

Note. All three columns present linear regression models using immediate refund as the dependent variable and the race and gender of the tester and the store agent as independent variables. Model 1 includes an interaction term between the race of the tester and the median prices at the store. Model 2 includes an interaction term between the race of the tester and a dummy variable which equals “1” if the store is a public corporation. Model 3 includes an interaction term between the race of the tester and the store’s annual revenues (in millions of dollars). All three models include parent company, day-of-audit, and time-of-audit fixed effects. Standard errors are presented in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

3. *Robustness Checks: Racial Disparities in Eventual Return Outcomes*

As detailed previously, if testers were initially denied an immediate refund, they were directed to ask to speak with the store's manager and to request a refund once more, regardless of whether they spoke with a manager. This Section discusses the final outcomes of these return attempts, which are categorized based on the results achieved after testers who had been initially refused a refund made a subsequent refund request. For those testers who received an immediate cash refund, both their initial and subsequent return outcomes were classified as "refund" and included in the analysis. Regression Table 3 presents the results.

REGRESSION TABLE 3: EVENTUAL RETURN OUTCOMES BY TESTER RACE AND GENDER AND OTHER VARIABLES

Eventual Refund as Dependent Variable				
	(1)	(2)	(3) + Agent Demographics	(4) + Interactions
White tester	0.0814* (0.0425)	0.0746 (0.0577)	0.0828* (0.0428)	0.0988 (0.107)
Female tester	0.0251 (0.0478)	0.0175 (0.0643)	0.0191 (0.0486)	0.0435 (0.114)
White tester # Female tester		0.0160 (0.0906)		-0.00598 (0.152)
White agent			0.0190 (0.0525)	0.0786 (0.118)
Male agent			0.0453 (0.0566)	0.101 (0.152)
BF tester # BM agent				0.0442 (0.216)
BM tester # WM agent				-0.167 (0.209)
BF tester # WF agent				-0.0995 (0.170)
BF tester # WM agent				-0.182 (0.259)
WM tester # BF agent				-0.0529 (0.194)

WM tester # BM agent	-0.163 (0.198)			
WF tester # BF agent	-0.129 (0.167)			
WM tester # WM agent	-0.0741 (0.267)			
WF tester # WF agent	-0.0787 (0.157)			
WF tester # WM agent	-0.0512 (0.247)			
Constant	0.0352 (0.0381)	0.0381 (0.0415)	0.0142 (0.0513)	-0.00418 (0.0778)
Observations	197	197	197	197
R²	0.434	0.434	0.438	0.454

Note. Eventual refund offers are used in the first four models as the dependent variable. Model 1 regresses the dependent variable on the race and gender of the tester. Model 2 adds an interaction term between the race of the tester and the gender of the tester. Model 3 adds the race and gender of the last store representative who interacted with the tester, whether it was a store clerk or manager (as perceived and reported by the tester) as controls. Model 4 adds a four-way interaction between the demographics (race and gender) of the tester and those of the store representative (as perceived and reported by the tester). All models include store, parent company, day-of-audit, and time-of-audit fixed effects. Standard errors are presented in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

NORTHWESTERN UNIVERSITY LAW REVIEW

REGRESSION TABLE 3 (CONT.): EVENTUAL RETURN OUTCOMES BY TESTER RACE AND GENDER
AND OTHER VARIABLES

Eventual Rejection as Dependent Variable				
	(5)	(6)	(7) + Agent Demographics	(8) + Interactions
White tester	−0.252*** (0.0644)	−0.328*** (0.0865)	−0.253*** (0.0650)	−0.537*** (0.155)
Female tester	−0.0719 (0.0744)	−0.158 (0.0991)	−0.0776 (0.0759)	−0.281 (0.174)
White tester # Female tester		0.181 (0.139)		0.375* (0.226)
White agent			−0.0113 (0.0807)	−0.150 (0.174)
Male agent			0.0706 (0.0864)	−0.179 (0.222)
BF tester # BM agent				0.0962 (0.316)
BM tester # WM agent				0.342 (0.311)
BF tester # WF agent				0.279 (0.252)
BF tester # WM agent				0.370 (0.381)
WM tester # BF agent				0.488* (0.282)
WM tester # BM agent				0.473 (0.289)
WF tester # BF agent				0.303 (0.250)
WM tester # WM agent				0.538 (0.392)
WF tester # WF agent				0.220 (0.237)
WF tester # WM agent				0.196 (0.361)

Constant	0.429*** (0.0582)	0.460*** (0.0628)	0.447*** (0.0778)	0.544*** (0.114)
Observations	186	186	186	186
R²	0.473	0.481	0.484	0.522

Note. Models 5 to 8 consist of the same four linear regressions as models 1 to 4, only this time using eventual rejections (i.e., refusals to offer refund, store credit, or an exchange) as the dependent variable. All models include store, parent company, day-of-audit, and time-of-audit fixed effects. Standard errors are presented in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

As Regression Table 3 illustrates, controlling for gender, white testers were about 8 percentage points more likely to be eventually offered a refund compared to Black testers.¹¹⁴ White testers were, controlling for gender, about 25 percentage points less likely to experience eventual rejection than were Black testers.¹¹⁵

Notably, there was no significant interaction between seller and tester demographics (race and gender) in terms of eventual refund offers. There was, however, a significant interaction between the demographics of the store agent (whether it was the store clerk or the manager) and the tester in the context of eventual rejections. Specifically, the racial disparities favoring white men in terms of the likelihood of facing rejection were less pronounced when testers engaged with Black male store agents. This interaction was the sole significant demographic interplay observed affecting the outcomes of customer interactions with store agents.¹¹⁶ The nuanced dynamics between the demographics of consumers and store agents merit further exploration to better understand these intricate relationships and their implications.

Figures 3A and 3B visualize these findings. Figure 3A displays the proportions of Black and white testers who—controlling for gender—eventually received a refund or were denied a return. Figure 3B further disaggregates the results by race and gender, capturing all four race–gender subgroups.

Figure 3A reveals that racial disparities in treatment not only persisted but also widened in terms of the likelihood of complete rejection. The racial gap increased from 14.5 percentage points at the immediate stage to 25.2 percentage points at the eventual stage. This widening disparity underscores the idea that, although requesting managerial intervention can

¹¹⁴ This difference was significant at the 10% level (see column (1)) and remained significant and similar in magnitude when controlling for the store agent's demographics (see column (3)).

¹¹⁵ This difference was significant at the 1% level (see column (5)) and remained significant and similar in magnitude when controlling for the store agent's demographics (see column (7)).

¹¹⁶ This difference was significant at the 10% level (see column (8)).

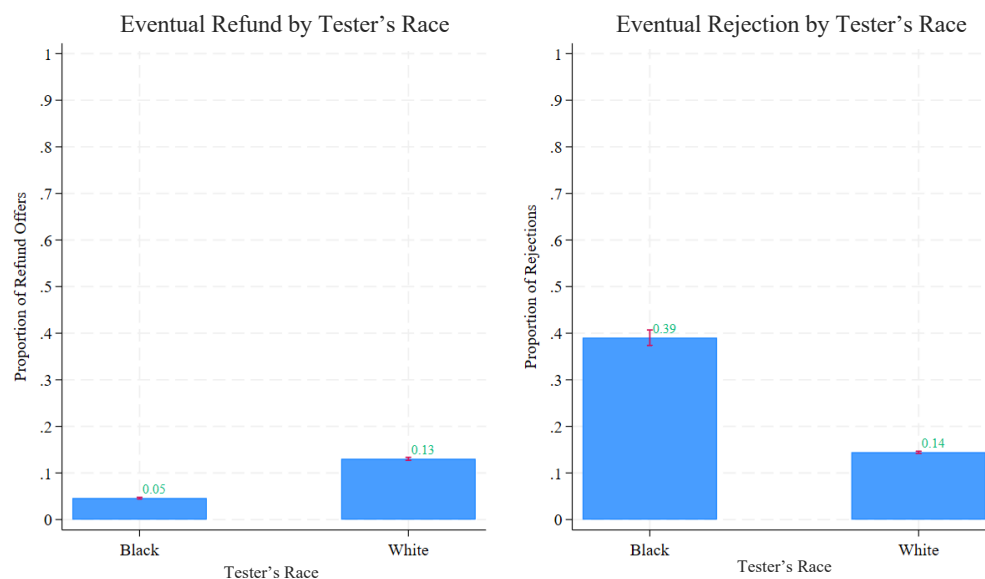
enhance the chances of receiving more accommodating treatment, such strategies fall short of bridging the racial divide in service experiences. Although behavioral science suggests that involving a manager might shift decision-making into a more deliberative mode—potentially reducing racial bias—this study finds the opposite: the racial gap actually grew wider. This suggests that managerial intervention, rather than mitigating bias, may inadvertently amplify it by adding another layer of discretion, allowing preexisting biases to potentially influence outcomes even more.

As Figure 3B illustrates, Black male testers experienced the worst return outcomes—both in terms of the lowest refund rates and the highest rejection rates. Black male testers faced a 46% rejection rate, nearly three times higher than white female testers (16%) and more than three times higher than white male testers (13%). These findings highlight the compounded disadvantage faced by Black men in this context, where both race and gender intersect to produce the most unfavorable outcomes.

This pattern aligns with intersectionality theory, which emphasizes that discrimination is not only shaped by race or gender alone but by their combined effects.¹¹⁷ Here, Black male testers appear to face a unique form of compounded disadvantage, suggesting that biases in consumer interactions may operate at the intersection of both racial and gendered perceptions.

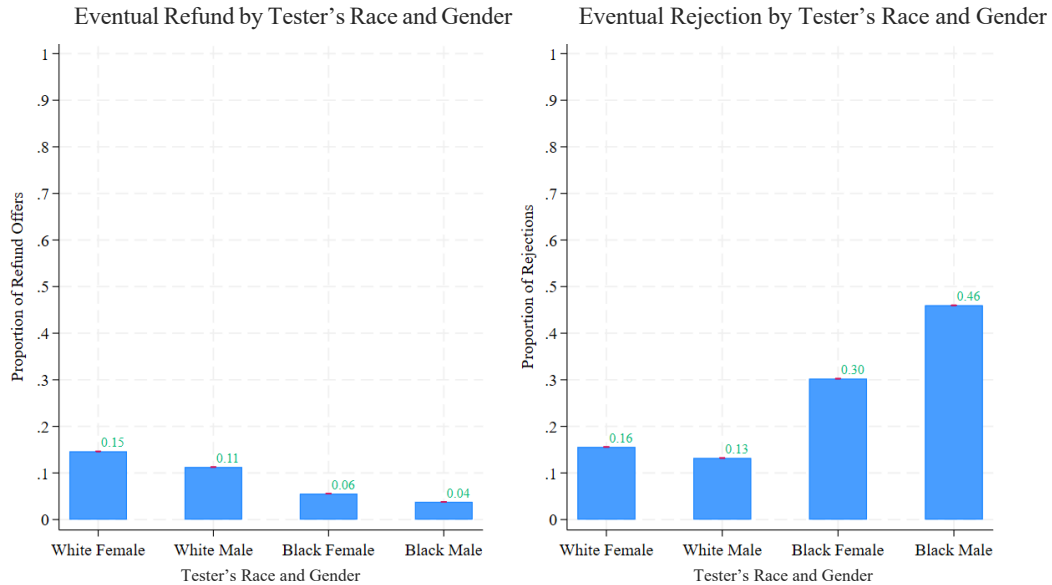
¹¹⁷ See *infra* notes 126–127 and accompanying text.

FIGURE 3A: EVENTUAL RETURN OUTCOMES BY TESTER'S RACE



Note. Figure 3A presents the predicted probabilities of eventual refund offers and rejections, disaggregated by tester race. The estimates are based on a regression model that controls for tester gender and includes fixed effects for store, parent company, day of audit, and time of audit (columns (1) and (5) in Regression Table 3). The bars represent the adjusted mean probabilities of refund offers and rejections for Black and white testers, with error bars indicating 95% confidence intervals.

FIGURE 3B: EVENTUAL RETURN OUTCOMES BY TESTER'S RACE AND GENDER



Note. Figure 3B presents the predicted probabilities of eventual refund offers and rejections, disaggregated by tester race and gender. The estimates are based on a regression model that controls for tester gender and includes fixed effects for store, parent company, day of audit, and time of audit (columns (2) and (6) in Regression Table 3). The bars represent the adjusted mean probabilities of refund offers and rejections for Black and white testers, with error bars indicating 95% confidence intervals.

4. *Is It Really About Discretion?*

This study sheds light on the discrepancies in the enforcement of retail return policies, revealing a pattern where deviations from the formal return policies disproportionately benefit white customers compared to similarly situated Black customers. This pattern prompts a critical inquiry: might the extent of racial disparities diminish in settings where store agents have limited or no discretion to deviate from official return policies?

To investigate this question, the study leverages natural variation in return policies across different retailers. While all participating stores required a receipt for refund, they varied in the degree of discretion granted to store clerks for issuing store credit or exchanges without a receipt. Approximately 63.5% of the audited stores in the study explicitly permitted their salesclerks to offer store credit or exchanges in the absence of a receipt, whereas 36.5% did not formally authorize such practices. This variation in

policy provided a unique opportunity to examine whether racial biases—manifested through the disproportionate offering of “concessions” (i.e., store credit or exchanges) to white customers compared to Black customers in similar situations—were more pronounced in settings where clerks explicitly had more discretionary power.

The analysis, detailed in Regression Table 4, categorizes stores based on their policy regarding discretionary authority to offer concessions (i.e., exchanges or store credit) into two groups: “discretion” and “no discretion.” It then examines racial disparities in both immediate and eventual concession offerings within each category.

Regression Table 4 includes four linear regression models. Models 1 and 3 use the “discretion” group of stores, while models 2 and 4 use the “no discretion” group. In models 1 and 2, the dependent variable is “immediate concession,” a binary variable coded as “1” if an immediate concession (refund, store credit, or exchange) was offered, and “0” otherwise. In models 3 and 4 the dependent variable is an eventual concession, a binary variable coded as “1” if a concession was offered either immediately or after the tester requested to speak with management, and “0” otherwise. All models include controls for the gender of the tester and the (perceived) race and gender of the store representative. All models also include store, parent company, day-of-audit, and time-of-audit fixed effects.

REGRESSION TABLE 4: CONCESSIONS BY STORE TYPE (“DISCRETION” VERSUS “NO-DISCRETION” STORES)

Immediate Refund as Dependent Variable				
	(1) Discretion: Immediate Concessions	(2) No Discretion: Immediate Concessions	(3) Discretion: Eventual Concessions	(4) No Discretion: Eventual Concessions
White tester	0.185** (0.0831)	0.0368 (0.122)	0.257*** (0.0704)	0.221** (0.0977)
Female tester	0.193** (0.0853)	0.0225 (0.160)	0.0649 (0.0749)	0.133 (0.129)
White agent	0.0607 (0.107)	0.188 (0.148)	0.127 (0.0893)	-0.0448 (0.119)
Male agent	0.0816 (0.104)	0.293* (0.163)	0.0126 (0.0893)	-0.0996 (0.152)
Constant	0.478*** (0.0848)	0.358** (0.142)	0.537*** (0.0773)	0.588*** (0.115)
Observations	128	67	128	67
R²	0.539	0.516	0.598	0.628

Note. This table includes four linear regression models. In models 1 and 2, the dependent variable is “immediate concession,” a dummy variable which equals “1” if an immediate concession (refund, store credit, or exchange) was offered, “0” otherwise. In models 3 and 4 the dependent variable is an eventual concession, a dummy variable which equals “1” if a concession was offered either immediately or after the tester requested to speak with management, “0” otherwise. Models 1 and 3 include only the subset of stores with a formal discretionary authority to offer concessions, while models 2 and 4 include only the subset of stores with no such discretionary language in their formal policies. All models control for the gender of the tester and the (perceived) race and gender of the store representative. All models also include store, parent company, day-of-audit, and time-of-audit fixed effects. Standard errors are presented in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The findings reveal a stark racial disparity in stores that grant their staff explicit discretionary powers: white customers experienced a significantly greater likelihood—18.5 percentage points higher—of obtaining immediate concessions compared to Black customers.¹¹⁸ Conversely, in stores lacking such discretionary policies, there is no significant racial benefit.¹¹⁹ These results suggest a potential association between discretionary practices and discriminatory outcomes, underscoring the critical role that such policies may play in driving the racial discrepancies observed.

¹¹⁸ See column (1). This racial difference is significant at the 5% level.

¹¹⁹ See column (2). White testers were 3 percentage points more likely to receive a concession, but this difference was not statistically significant.

However, the interpretation of these findings demands caution. The identified correlation between discretionary policies and racial disparities does not establish causality. Differences inherent to stores with and without such policies might also contribute to the observed disparities. Additionally, the gap between the two types of stores narrows when considering eventual concessions—those granted after an initial denial and subsequent escalation to management—with Black customers between 22 and 25.7 percentage points less likely than white customers to be offered concessions.¹²⁰ This pattern implies that managerial discretion, irrespective of formal policy language, could still exhibit a bias favoring white customers over Black customers.

D. Discussion

This investigation revealed a concerning pattern in retail settings: Salesclerks inconsistently enforce return policies, disproportionately disadvantaging Black consumers. This practice aligns with emerging research indicating that sellers often empower their employees with discretionary authority under the guise of standardized policies.¹²¹ In fact, these findings suggest that said discretion might be exercised in a racially biased manner, undercutting the presumed uniformity of sellers' policies.

Despite retailers' official requirement for receipts in return transactions, this analysis identifies significant disparities: white customers, particularly white women, were more likely to receive offers for refunds, store credit, or exchanges than their Black counterparts, particularly Black men, under similar circumstances. The observed racial disparities in treatment not only persisted but expanded in scenarios involving negotiations with management following initial refusals, indicating a systemic advantage for white consumers at various transactional junctures.

Finally, this study elucidates the role of *discretion* in exacerbating racial disparities in contractual performance. Retail outlets with policies that formally allow discretionary returns witnessed more pronounced racial disparities, suggesting that official permission to deviate amplifies racial bias in decision-making.

Taken together, these findings contribute to the growing body of research on “shopping while Black,” emphasizing the nuanced forms of racial discrimination present in retail environments.¹²² They corroborate

¹²⁰ See columns (3) and (4).

¹²¹ See *supra* Section I.B.

¹²² See *supra* Section I.A.

earlier studies showing that Black shoppers, and especially Black men, are subject to harsher treatment and more surveillance in retail settings,¹²³ while white shoppers, especially white women,¹²⁴ enjoy a distinct level of privilege.¹²⁵

The stark differential treatment between white women and Black men underscores the critical need for an intersectional approach in studying discrimination, acknowledging the complex interplay of race and gender.¹²⁶ Originally conceived to spotlight the unique challenges Black women face,¹²⁷ intersectionality is now seen through this study's findings to warrant broader

¹²³ See, e.g., Schreer et al., *supra* note 18, at 1437–40 (finding that Black men are treated with greater suspicion by retail security guards than are any other race–gender subgroup members).

¹²⁴ See, e.g., WILLIAMS, *supra* note 30, at 106, 131 (noting a similar bias towards white female shoppers in toy stores). This bias toward white women may stem from stereotypes that paint them as more credible and trustworthy or as the primary decision-makers in household retail consumption. See, e.g., PATRICIA HUDDLESTON & STELLA MINAHAN, CONSUMER BEHAVIOR: WOMEN AND SHOPPING 1–8 (2011) (observing that “[w]omen matter because, worldwide, they control so much of consumer spending,” and because “[i]n most cases, women are still expected to maintain the household and remain responsible for domestic chores (including shopping)”). Additionally, societal perceptions of white women’s sense of entitlement could also play a role. See, e.g., Apryl Williams, *Black Memes Matter: #LivingWhileBlack with Becky and Karen*, SOC. MEDIA + SOC’Y, Oct.–Dec. 2020, at 1, 2, 10 (observing that the “Becky” and “Karen” memes have become a symbol of white women’s sense of entitlement); see also Terri D. Conley, *Beautiful, Self-Absorbed, and Shallow: People of Color Perceive White Women as an Ethnically Marked Category*, 43 J. APPLIED SOC. PSYCH. 45, 47–48, 53 (2012) (making similar observations).

¹²⁵ A recent Australian audit study has revealed an analogous tendency to bend rules to accommodate white passengers more than Black passengers. See Redzo Mujcic & Paul Frijters, *Still Not Allowed on the Bus: It Matters If You’re Black or White!* (Institute for the Study of Labor, IZA DP No. 7300, 2013), <https://docs.iza.org/dp7300.pdf> [<https://perma.cc/EF7G-2U26>] (finding that bus drivers in Brisbane, Australia were nearly twice as willing to let white testers ride free than they were for Black testers similarly explaining that they did not have enough money but needed to get to a specific station); see also Ian Ayres, Opinion, *When Whites Get a Free Pass*, N.Y. TIMES (Feb. 24, 2015), http://www.nyt.com/2015/02/24/opinion/research-shows-white-privilege-is-real.html?referrer=&_r=2 [<https://perma.cc/TV4Y-MZHN>] (describing this and other empirical findings of “white privilege” and racial discrimination in public places).

¹²⁶ For prominent recent scholarship advocating for the adoption of an intersectional approach, which emphasizes how racial and gender identities interact, in discrimination research, see, for example, Sumi Cho, Kimberlé W. Crenshaw & Leslie McCall, *Toward a Field of Intersectionality Studies: Theory, Applications, and Praxis*, 38 SIGNS: J. WOMEN CULTURE & SOC’Y 785, 787 (2013), which comments that intersectionality has “played a major role in facilitating consideration of gender, race, and other axes of power in a wide range of political discussions and academic disciplines”; and Catharine A. MacKinnon, *Intersectionality as Method: A Note*, 38 SIGNS: J. WOMEN CULTURE & SOC’Y 1019, 1020 (2013), which notes that intersectionality “adds the specificity of sex and gender to race and ethnicity, and racial and ethnic specificity to sex and gender.”

¹²⁷ See Devon W. Carbado, Kimberlé Williams Crenshaw, Vickie M. Mays & Barbara Tomlinson, *Intersectionality: Mapping the Movements of a Theory*, 10 DUBOIS REV. 303, 303 (2013) (observing that the term “intersectionality” has been introduced “to address the marginalization of Black women”).

application, suggesting it should also encompass the complex and varied forms of discrimination faced by Black men.¹²⁸

The intersectionality of race and class could also play a role in store agents' decisions.¹²⁹ While the subject study attempted to control for class and socioeconomic status by instructing all testers to wear similar clothing, it must be acknowledged that the same clothing (in this case, jeans and a t-shirt) may convey different meanings when worn by members of different racial or ethnic groups.¹³⁰ Experimental evidence confirms this hypothesis, suggesting that clothing perceived as "casual" when worn by white men—e.g., baggy jeans, a sweatshirt, and a baseball cap—can be perceived as "threatening" when worn by Black men,¹³¹ and consequently lead to racial bias in decision-making.¹³² Thus, to the extent that people conflate race with class, such inferences may have played a role in the store agents' decisions.¹³³

E. Implications and Generalizability

The "new" form of retail race discrimination uncovered in this Article risks undermining the mental well-being, life satisfaction, and self-esteem of minority consumers.¹³⁴ Moreover, sellers' strategy of authorizing on-

¹²⁸ Indeed, some commentators have argued that the intersectionality paradigm has already shifted to engage Black men. *See id.* at 310.

¹²⁹ On the important role of the intersection between race and class in shaping perceptions and interpersonal behavior, see Cecilia L. Ridgeway & Tamar Kricheli-Katz, *Intersecting Cultural Beliefs in Social Relations: Gender, Race, and Class Binds and Freedoms*, 27 GENDER & SOC'Y 294, 304 (2013).

¹³⁰ Cf. Sen & Wasow, *supra* note 10, at 500 (proposing that "in experimental or empirical contexts, race should be understood as a composite variable or 'bundle of sticks,'" allowing "race to be disaggregated into constitutive elements, some of which can be manipulated experimentally").

¹³¹ *See* M. Kimberly MacLin & Vivian Herrera, *The Criminal Stereotype*, 8 N. AM. J. PSYCH. 197, 203–04 (2006); Kimberly Barsamian Kahn & Paul G. Davies, *What Influences Shooter Bias? The Effects of Suspect Race, Neighborhood, and Clothing on Decisions to Shoot*, 73 J. SOC. ISSUES 723, 734, 736 (2017) (finding racial bias against Black suspects when they wear casual clothing, consisting of baggy gray sweatshirt, a gray headband, and a black baseball cap worn to the side).

¹³² Kahn & Davies, *supra* note 131, at 734, 736 (finding significant racial bias in decisions to "shoot" suspects when the suspects—Black and white individuals—wear casual clothing but not when they wear a more professional attire).

¹³³ *See, e.g.*, Jonathan B. Freeman, Andrew M. Penner, Aliya Saperstein, Matthias Scheutz & Nalini Ambady, *Looking the Part: Social Status Cues Shape Race Perception*, PLOS ONE, Sept. 26, 2011, at 1, 2, 7 (finding that social and contextual factors guide the perception of race).

¹³⁴ *See* Yin Paradies, Jehonathan Ben, Nida Denson, Amanuel Elias, Naomi Priest, Alex Pieterse, Arpana Gupta, Margaret Kelaher & Gilbert Gee, *Racism as a Determinant of Health: A Systematic Review and Meta-Analysis*, PLOS ONE, Sept. 23, 2015, at 1, 10–11 (surveying evidence that racial discrimination is associated with negative health consequences); Michael T. Schmitt, Nyla R. Branscombe, Tom Postmes & Amber Garcia, *The Consequences of Perceived Discrimination for Psychological Well-Being: A Meta-Analytic Review*, 140 PSYCH. BULL. 921, 922 (2014) (finding a significant negative relationship between

the-ground discretion in the exercise of their contracts and policies may very well extend beyond the retail industry, suggesting that discretionary contract enforcement could be a widespread source of racial discrimination across consumer markets. Indeed, preliminary evidence indicates similar biases in various sectors, from housing and mortgage services, to insurance and healthcare, hinting at a pervasive problem.¹³⁵

The implications are profound: discriminatory contract enforcement not only inflicts direct harm on racial minorities by providing them with inferior goods or services but also perpetuates socioeconomic disparities through a hidden form of price discrimination.¹³⁶ In this context, “price discrimination” refers to minority customers paying the same price as others but receiving lower quality goods or services in return. This creates an imbalance where minority consumers unknowingly subsidize the cost of “extras” or better treatment that more privileged groups receive—such as waived fees, flexible return policies, or other favorable adjustments. Effectively, businesses factor these “freebies” into their pricing model, passing the cost onto groups perceived as “less desirable” customers.¹³⁷ In this context, price

perceived discrimination and mental health and a somewhat weaker, but still significant, negative association with self-esteem and life satisfaction). For a general assessment of the social costs produced by marketplace race discrimination, see Peter Siegelman, *Racial Discrimination in “Everyday” Commercial Transactions: What Do We Know, What Do We Need to Know, and How Can We Find Out?*, in A NATIONAL REPORT CARD ON DISCRIMINATION IN AMERICA: THE ROLE OF TESTING 69 (Michael Fix & Margery Austin Turner eds., 1998).

¹³⁵ For example, informal interviews with store clerks in the Chicago area conducted in preparation for this study suggest that white customers are disproportionately granted more delivery-fee waivers compared to similarly situated Black customers. See *infra* Part IV. Similarly, evidence suggests that corporate landlords disproportionately exercise their contractual right to evict Black tenants, as compared to similarly situated white tenants. See, e.g., Chris Arnold, *Corporate Landlord Evicts Black Renters at far Higher Rates than Whites, Report Finds*, NPR (June 3, 2021, 5:01 AM), <https://www.npr.org/2021/06/03/1001404416/corporate-landlord-evicts-black-renters-at-far-higher-rates-than-whites-report-f> [<https://perma.cc/WEF4-6YWH>]. Furthermore, a 2021 survey reveals evidence of discriminatory practices by insurance companies in their handling of insurance claims from Black homeowners. See Emily Flitter, *New Suit Uses Data to Back Racial Bias Claims Against State Farm*, N.Y. TIMES (Dec. 14, 2022), <https://www.nyt.com/2022/12/14/business/state-farm-racial-bias-lawsuit.html> [<https://perma.cc/KEU5-4V34>]. This evidence underlines the potential presence of similar disparities in other sectors.

¹³⁶ Indeed, in real life, other status signals, such as a consumer’s brand-name or expensive attire, may bolster or attenuate observed racial and gender differences in treatment. See, e.g., Freeman et al., *supra* note 133 (finding that study participants presented with photos of individuals with racially ambiguous faces were more likely to categorize those individuals as “Black” when they wore “low-status” attire).

¹³⁷ This study has focused on discrimination against Black Americans. Future studies could expand this line of research by investigating whether and to what extent other racial and ethnic minorities, including Hispanics, Asians, and Native Americans, are similarly subject to discriminatory enforcement of consumer contracts. Researchers could also test whether contracts are enforced differently within the

discrimination does not refer to charging different prices based on customer demographics (as it often does in traditional economic models). Instead, it highlights how minority consumers are charged the same price as others but are given inferior products or services, leading to a hidden financial burden. Essentially, they are paying more for less, while more privileged customers benefit from extra perks that are effectively built into the price of the goods or services but are disproportionately unavailable to minorities.

Addressing this issue requires education—so that lawmakers and the public become aware of the problem—as well as actionable legal strategies to mitigate the discrimination in consumer markets. Indeed, this situation highlights the need for a critical reassessment of current legal frameworks to better protect consumers from racial bias and its cascading socioeconomic effects. Of course, before one can devise effective strategies, it is important to understand the root causes behind the documented discrepancies in contractual enforcement. The next Part of this Article explores those potential explanations, with the goal of laying the groundwork for informed legal and policy interventions.

IV. POTENTIAL EXPLANATIONS

A. Psychological Research on Implicit Bias

What could explain the racial disparities observed in salespeople's treatment of consumers' return requests? For many years, the most widely accepted explanation for discriminatory outcomes was bigotry or animosity toward “out-group” members—and, conversely, better treatment of “in-group” members.¹³⁸ In fact, some of the observed disparities in the study can be explained this way. Specifically, the finding that white male customers experience significantly greater favoritism when interacting with white store clerks compared to nonwhite store clerks resonates with

Black population based on different skin tones. Although colorism and racism are often intertwined, it is entirely possible that people belonging to the same race might suffer from different degrees (or forms) of discrimination based on their skin tones. Indeed, empirical research has shown that, at least in the employment context, darker skin tones have a negative impact on employment possibilities and outcomes. See, e.g., Tamar Kricheli-Katz, Tali Regev, Shay Lavie, Haggai Porat & Ronen Avraham, *Those Who Tan and Those Who Don't: A Natural Experiment on Colorism*, PLoS ONE, July 24, 2020, at 1, 2 (2020).

¹³⁸ See, e.g., Nat'l Rsch. Council of the Nat'l Acad., *Theories of Discrimination*, in MEASURING RACIAL DISCRIMINATION 56 (Rebecca M. Blank, Marilyn Dabady & Constance F. Citro eds., 2004) (“Most people’s concept of racial discrimination involves explicit, direct hostility expressed by whites toward members of a disadvantaged racial group.”).

established research on in-group bias.¹³⁹ Numerous studies confirm this phenomenon and suggest that individuals' group affiliations subtly influence their perceptions and behaviors, often unconsciously,¹⁴⁰ frequently resulting in leniency towards in-group members and harsher treatment of those from out-groups.¹⁴¹ Such in-group biases have been identified in various fields, including sports and the judicial system.¹⁴²

Recent empirical research in other domains, however, suggests that racial disparities in treatment often stem from implicit, unconscious racial bias.¹⁴³ Research indicates that even the most well-intentioned decision-makers may possess unconscious or implicit biases against certain groups.¹⁴⁴

¹³⁹ See, e.g., C. Bram Cadsby, Ninghua Du & Fei Song, "In-Group" Favoritism and Moral Decision-Making, 128 J. ECON. BEHAV. & ORG. 59, 60 (2016) (presenting evidence that even without direct personal benefit, people will cheat to protect in-group members at the expense of out-group members). Several theories attempt to explain the occurrence of in-group bias; however, the social identity theory proposed by Polish social psychologist Henri Tajfel and colleagues is perhaps the most cited. This theory is rooted in the human propensity to categorize. According to this theory, this cognitive process, which exaggerates differences between in-group and out-group members, serves essential functions such as quickly identifying allies or threats, and operates unconsciously, beyond the decision-maker's intent or awareness. See Henri Tajfel, *Social Psychology of Intergroup Relations*, 33 ANN. REV. PSYCH. 1, 2 (1982); Michael Billig & Henri Tajfel, *Social Categorization and Similarity in Intergroup Behaviour*, 3 EUR. J. SOC. PSYCH. 27, 47–49 (1973).

¹⁴⁰ See, e.g., Billig & Tajfel, *supra* note 139, at 47–48 (finding that participants displayed in-group bias even when their group was formed based on random assignment).

¹⁴¹ See, e.g., *id.*

¹⁴² On bias in sports, see generally Daniel L. Wann & Frederick G. Grieve, *Biased Evaluations of In-Group and Out-Group Spectator Behavior at Sporting Events: The Importance of Team Identification and Threats to Social Identity*, 145 J. SOC. PSYCH. 531 (2005), which reports on a study in which a team's fans evaluated the conduct of both their own team and the opposing team's fans after a baseball game and found a pronounced in-group bias. For the judicial system, see generally Oren Gazal-Ayal & Raanan Sulitzeanu-Kenan, *Let My People Go: Ethnic "In-Group" Bias in Judicial Decisions: Evidence from a Randomized Natural Experiment*, 7 J. EMPIRICAL LEGAL STUD. 403 (2010), which found, based on a randomized natural experiment in the Israeli judiciary, that Jewish judges displayed in-group ethnic bias in detention decisions; and David S. Abrams, Marianne Bertrand & Sendhil Mullainathan, *Do Judges Vary in Their Treatment of Race?*, 41 J. LEGAL STUD. 347 (2012), which found significant racial disparities in favor of white defendants in sentencing decisions rendered by white judges in Cook County, Illinois.

¹⁴³ For a discussion of some of the psychological and social science research on implicit bias, see, for example, Anthony G. Greenwald & Linda H. Krieger, *Implicit Bias: Scientific Foundations*, 94 CALIF. L. REV. 945 (2006); Christine Jolls & Cass R. Sunstein, *The Law of Implicit Bias*, 94 CALIF. L. REV. 969 (2006); Jerry Kang, Judge Mark Bennett, Devon Carbado, Pam Casey, Nilanjana Dasgupta, David Faigman, Rachel Godsil, Anthony G. Greenwald, Justin Levinson & Jennifer Mnookin, *Implicit Bias in the Courtroom*, 59 UCLA L. REV. 1124 (2012); and Erik J. Girvan, *On Using the Psychological Science of Implicit Bias to Advance Anti-Discrimination Law*, 26 GEO. MASON U. C.R.L.J. 1 (2015).

¹⁴⁴ MAHZARIN R. BANAJI & ANTHONY G. GREENWALD, *BLINDSPOT: HIDDEN BIASES OF GOOD PEOPLE* xii–xv (2016); see also Nicholas Kristof, Opinion, *Is Everyone a Little Bit Racist?*, N.Y. TIMES (Aug. 27, 2014), <https://www.nytimes.com/2014/08/28/opinion/nicholas-kristof-is-everyone-a-little-bit->

These implicit biases stem from the automatic ways in which people process and categorize information and can be difficult to uncover.¹⁴⁵ Indeed, some observers contend that unconscious bias now constitutes the most prevalent form of discrimination in contemporary society.¹⁴⁶

This shift towards understanding implicit biases as a source of discriminatory practices marks a critical evolution in our comprehension of prejudice.¹⁴⁷ It suggests that disparities in treatment, such as those observed in this study—where Black consumers faced less favorable outcomes in return requests—may not stem from a conscious desire to discriminate, but rather from ingrained stereotypes associating race with differences in consumers' expectations, perceptions, and behaviors.¹⁴⁸ Implicit bias literature, particularly findings from the Implicit Association Test (IAT),¹⁴⁹ indicates that biases tend to be more pronounced in fast-paced, immediate decision-making environments.¹⁵⁰ In this study, interactions with store clerks might involve quick judgments, where implicit biases are likely to surface. However, managerial discretion may involve slower, more deliberate

racist.html [https://perma.cc/PNP7-C49Q] (describing the recent problem of people who “consider themselves enlightened, who intellectually believe in racial equality, who deplore discrimination,” yet who still harbor unconscious biases).

¹⁴⁵ See, e.g., Jolls & Sunstein, *supra* note 143, at 973 (“[I]mplicit bias—like many of the heuristics and biases emphasized elsewhere—tends to have an automatic character . . .” (emphasis added)); Linda Hamilton Krieger & Susan T. Fiske, *Behavioral Realism in Employment Discrimination Law: Implicit Bias and Disparate Treatment*, 94 CALIF. L. REV. 997, 1030–31 (2006); Kristin A. Lane, Jerry Kang & Mahzarin R. Banaji, *Implicit Social Cognition and Law*, 3 ANN. REV. L. & SOC. SCI. 427, 428–29 (2007); Jeffrey J. Rachlinski & Cynthia R. Farina, *Cognitive Psychology and Optimal Government Design*, 87 CORNELL L. REV. 549, 555–56 (2002); Greenwald & Krieger, *supra* note 143, at 950–52, 959–62.

¹⁴⁶ See, e.g., Linda H. Krieger, *The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity*, 47 STAN. L. REV. 1161, 1164 (1995) (observing that unconscious forms of bias “represent today’s most prevalent type of discrimination”); David B. Oppenheimer, *Negligent Discrimination*, 141 U. PA. L. REV. 899, 899 (1992) (“[R]ecent studies support the assertion that most discrimination is not the result of malice, hatred, ill will, or bigotry: it is the result of unintended and unconscious stereotyping.”); Charles R. Lawrence III, *The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism*, 39 STAN. L. REV. 317, 335 (1987) (making a similar observation).

¹⁴⁷ See, e.g., Jerry Kang & Kristin Lane, *Seeing Through Colorblindness: Implicit Bias and the Law*, 58 UCLA L. REV. 465, 519–20 (2010) (suggesting modern empirical evidence now rebuts “facile assumptions of colorblindness”).

¹⁴⁸ See, e.g., Greenwald & Krieger, *supra* note 143, at 950–52 (suggesting implicit biases might be especially problematic because they “produce behavior that diverges from a person’s avowed or endorsed beliefs or principles”).

¹⁴⁹ The IAT is the most widely used tool for detecting implicit bias. For descriptions of how the IAT works, see *id.* at 952–67.

¹⁵⁰ See, e.g., L. Song Richardson & Phillip A. Goff, *Implicit Racial Bias in Public Defender Triage*, 122 YALE L.J. 100, 105 (2013) (observing, and citing evidence, that implicit biases “thrive in situations where individuals make decisions quickly with imperfect information and when they are cognitively depleted, anxious, or distracted”).

decision-making, raising questions about whether such decisions are immune to bias or if the additional discretion might actually exacerbate existing stereotypes.

The findings suggest that managerial involvement does not necessarily reduce racial bias but may, in some cases, reinforce it, particularly if discretion is exercised without sufficient checks. This points to an opportunity for policy interventions: businesses could implement procedures that slow down the decision-making process, encouraging more deliberative reasoning. By designing return policies that require additional steps or structured guidelines—such as using data-driven decision-making tools or mandating written justifications for rejections (as further detailed below)—businesses could help mitigate the influence of unconscious biases and promote fairer outcomes for minority consumers.

In the context of retail returns, differing perceptions based on prevailing stereotypes—conscious or unconscious—may influence interactions between store clerks and customers in several ways. First, clerks may perceive white consumers—particularly white women—as having higher transactional expectations, believing these consumers are more likely to express dissatisfaction or damage the store’s reputation if their expectations are unmet. Second, clerks may associate white customers with greater purchasing power or higher socioeconomic status (SES), viewing them as more valuable to the business. Third, there is concern that clerks may stereotype Black customers—especially Black men—as more likely to misuse return policies or engage in shoplifting. Regarding the first explanation concerning differential perceptions of customer expectations, if clerks believe that minority customers are less likely to demand accommodation or cause reputational damage compared to their white counterparts, they might overlook the requests of Black customers.¹⁵¹

¹⁵¹ In the real world, minority consumers may indeed be less likely to complain if they believe that their chances of obtaining redress are low. This, in turn, can create a vicious cycle: if Black consumers do not feel entitled to complain or do not anticipate a successful outcome after complaining, they may be less likely to register a complaint, thereby becoming even less likely to receive redress. At the same time, white consumers, who might not suffer from the same sense of impotence, will continue to complain, disproportionately benefiting from preferential treatment from the seller. *See, e.g.,* Marsha L. Richins, *An Investigation of Consumers’ Attitudes Toward Complaining*, 9 ADVANCES CONSUMER RSCH. 502, 503, 505 (1982); Marsha L. Richins, *Negative-Word-of-Mouth by Dissatisfied Consumers*, 47 J. MKTG. 68, 75–76 (1983) (observing that the likelihood of consumers complaining depends to a large extent on the perceived responsiveness of the seller).

Regarding the second explanation about inferences of SES and value to the store, research indicates a common association between race and class,¹⁵² suggesting that these intertwined perceptions influence how clerks treat customers.¹⁵³ This bias could lead to preferential treatment for white customers, who are perceived as wealthier and therefore more valuable.

Notably, the audit study presented in this Article tried to minimize the potential influence of perceived class or SES by having participants wear similar attire. However, the interpretation of clothing can still vary with the wearer's race or ethnicity, potentially affecting clerks' and managers' decisions.¹⁵⁴ Real-life scenarios might further complicate this, as additional status indicators, such as brand-name clothing, could influence how customers are treated, potentially reinforcing racial disparities in how contracts and policies are applied.¹⁵⁵ And this type of selective enforcement of contracts can have regressive distributional outcomes, with poor consumers less likely than affluent consumers to enjoy favorable concessions or deviations from the language of a standard contract.¹⁵⁶

Regarding the third explanation—the potential misuse of return policies—some evidence suggests that clerks might be predisposed to

¹⁵² See, e.g., Ridgeway & Kricheli-Katz, *supra* note 129, at 294, 304 (surveying the evidence about the common association between race and class); Amy J.C. Cuddy, Susan T. Fiske, Peter Glick & Jun Xu, *A Model of (Often Mixed) Stereotype Content: Competence and Warmth Respectively Follow from Perceived Status and Competition*, 82 J. PERSONALITY & SOC. PSYCH. 878, 891–92 (2002) (finding that people closely associate between “White people” and “Middle-class people,” as well as between “Black people” and “Blue-collar workers”); Andrew M. Penner & Aliya Saperstein, *Engendering Racial Perceptions: An Intersectional Analysis of How Social Status Shapes Race*, 27 GENDER & SOC’Y 319, 320, 326, 331–32 (2013) (finding that people who experienced events such as unemployment, incarceration, or poverty during a certain period were more likely to be perceived by interviewers as Black and less likely to be seen as white, even if they were classified differently in earlier phases of the survey, highlighting how socioeconomic changes can influence racial categorization by others).

¹⁵³ See, e.g., Ridgeway & Kricheli-Katz, *supra* note 129, at 311–12; Freeman et al., *supra* note 133, at 1, 8 (finding that racial differences in treatment may be bolstered or attenuated by other status signals, such as a consumer’s brand name or expensive attire); Zachary W. Brewster, *Racialized Customer Service in Restaurants: A Quantitative Assessment of the Statistical Discrimination Explanatory Framework*, 82 SOCIO. INQUIRY 3, 5 (2012); Zachary W. Brewster, Jonathan R. Brauer & Michael Lynn, *Economic Motivations and Moral Controls Regulating Discrimination Against Black and Hispanic Diners*, 56 SOCIO. Q. 506, 517 (2015) (finding significant associations between servers’ beliefs that Blacks are low-value customers and their self-reported propensity to racially profile Black consumers by providing them lower quality service).

¹⁵⁴ See, e.g., Sen & Wasow, *supra* note 10, at 499, 506 (citing studies showing the “subtle changes in cues such as . . . clothing . . . resulted in significant differences in how race or ethnicity operated as treatments”); Freeman et al., *supra* note 133, at 1–2 (discussing how “initial [racial] categorizations rely only on minimal physical cues” including clothing and face shape).

¹⁵⁵ See, e.g., Freeman et al., *supra* note 133, at 1–2; Sen & Wasow, *supra* note 10, at 506.

¹⁵⁶ For a similar observation, see Padi, *supra* note 64, at 827 (suggesting that selective performance of contracts might “privilege sophisticated parties and worsen existing social inequalities”).

suspect theft or policy abuse, particularly from Black customers and especially Black men.¹⁵⁷ In one notable study, for example, when participants were asked to envision a “typical shoplifter,” they consistently described a young, Black male—even though local crime statistics indicated that the vast majority of shoplifters were white and female.¹⁵⁸ Such biases could be even more pronounced in areas such as downtown Chicago, areas which are predominantly white and where racial dynamics might amplify discrimination against Black consumers.¹⁵⁹

B. Insights Gleaned from Interviews with Store Clerks

To identify the potential reasons behind the observed racial disparities in salespeople’s treatment of consumer return requests, this study also draws on original interviews with ten retail store clerks, all self-identifying as white, from the Chicago area. These individuals, sourced via Craigslist, included both past and present clerks from various retail establishments in Chicago. The interviews, which were designed as semi-structured dialogues, focused on the clerks’ interactions with customers and their application of return policies, particularly whether they would be willing to deviate from store rules or guidelines.¹⁶⁰

The findings illuminate the clerks’ perspectives and decision-making processes. For example, they reveal that the clerks would often make discretionary decisions based on whether they believed it would help to cultivate a loyal clientele. Thus, clerks would sometimes issue refunds if they recognized the customer or item.¹⁶¹ As one clerk at a high-end clothing store explained: “It’s upscale, so [we] want to be very customer-friendly, so [we]

¹⁵⁷ See, e.g., Shaun L. Gabbidon & George E. Higgins, *Public Opinion on the Use of Consumer Racial Profiling to Identify Shoplifters: An Exploratory Study*, 36 CRIM. JUST. REV. 201, 202–03 (2011); John Rappaport, *Criminal Justice, Inc.*, 118 COLUM. L. REV. 2251, 2290 n.235 (2018); Schreer et al., *supra* note 18, at 1432–44.

¹⁵⁸ See Jo Ann L. Asquith & Dennis N. Bristow, *To Catch a Thief: A Pedagogical Study of Retail Shoplifting*, 75 J. EDUC. FOR BUS. 271, 272–73 (2000).

¹⁵⁹ See, e.g., Lee, *supra* note 36, at 355–56 (“[B]usiness owners in [w]hite neighborhoods may more overtly object to the presence of Black customers because they feel that Blacks simply do not belong . . .”).

¹⁶⁰ The full interview scripts and transcripts are on file with *Northwestern University Law Review*.

¹⁶¹ For instance, one clerk mentioned that they could process a refund if they recognized the customer or the item. As that clerk explained, “[i]f it’s within a few days, one of us would remember the customer and be able to process a refund. When we didn’t remember them, we could look at the item and see if it was from our store. If it was, we would refund them. If not, we would turn customers away and say we couldn’t reimburse them. But we really wanted to build a good customer base, so we tried to be as accommodating as possible.” Interview with Interviewee no. 1 (Sept. 3, 2018) (transcript on file with *Northwestern University Law Review*).

were much more bendable with the rules, and [we]’re an expensive store so [we] don’t want to lose a customer fighting about \$30 if they are going to spend a thousand dollars.”¹⁶² Or, as another interviewee at a high-end clothing store observed: “If customers are not happy, you can offer them discounts . . . We had a policy, yet we had to do whatever we needed to do to keep the customer happy. . . . Policy goes out the door when you have an unhappy customer. Policy goes out the door because you’re trying to keep that customer happy, so [you do] whatever you can do.”¹⁶³ Most interviewees underscored their motivation to please customers as a primary reason for policy deviations, including accommodating customers with no receipts.

Additionally, some interviewees mentioned that incentives, such as earning a commission, influenced them to offer concessions. For example, one interviewee, who used to work in a high-end rug store, explained how he might be willing to waive the normal \$100 delivery fee if the customer asked and if the person seemed “really interested” because he knew he would earn a commission and that his boss “wouldn’t care.”¹⁶⁴

Other interviewees noted that the discretionary power to significantly deviate from return policies resided solely with managers, not salesclerks, but explained that they believed managers were primarily concerned with keeping customers happy. As one interviewee explained, “you would have to get the manager [be]cause we didn’t have any authority to do anything unless we were presented with a receipt.”¹⁶⁵ According to that interviewee, although there was “no hard or fast rule” on what the managers did, the salesclerk believed that managers “were just trying to keep the customer happy.”¹⁶⁶

At the same time, interviewees explained that stores needed to maintain a balance between accommodating customers and preventing abuse of lenient policies. As one interviewee explained, while someone might legitimately change their mind about an item—like a Halloween costume— “[m]any times, people would shoplift from other stores and then attempt

¹⁶² Interview with Interviewee no. 9 (Oct. 22, 2018) (transcript on file with *Northwestern University Law Review*).

¹⁶³ Interview with Interviewee no. 10 (Oct. 22, 2018) (alteration in original) (transcript on file with *Northwestern University Law Review*).

¹⁶⁴ Interview with Interviewee no. 3 (Sept. 12, 2018) (transcript on file with *Northwestern University Law Review*).

¹⁶⁵ Interview with Interviewee no. 2 (Sept. 21, 2018) (alteration in original) (transcript on file with *Northwestern University Law Review*).

¹⁶⁶ *Id.*

to return it [at her store], so [the store] had a strict return policy.”¹⁶⁷ She did note, however, that if she “saw someone with a kid” or someone she “recognized from the neighborhood,” she would make an exception. Otherwise, “no non-receipted returns were accepted.”¹⁶⁸

Other interviewees also talked about the receipt rule as a useful way to deal with customers who would try to return items that had clearly been used or damaged. As one interviewee explained, “People would return clothes that ha[d] perspiration stains on them. The first thing we would say is that it appears to have been used, or not even that, we would say that we can’t take it back without a receipt.”¹⁶⁹

The interviews confirmed that store agents often have discretion to deviate from stores’ formal policies, including return policies. Even more important, however, the interviews revealed a complex web of assumptions about race, gender, class, and customer behavior and suggest a relationship between these assumptions and employees’ decisions to deviate from the formal policies. Although the interviewed salesclerks vehemently denied that racial or gender bias influenced their deviation decisions, their nuanced responses suggest that demographic factors subtly influenced how they treated customers, potentially explaining some of the disparities in the application of return policies observed in the audit study.

In particular, the interviews uncovered a prevalent perception among clerks that white customers, and especially white women, have higher service expectations and that they are also more prone to voice their dissatisfaction if those expectations are not met. For example, the interviewee who used to work at a high-end rug store noted a pattern where fee waivers were predominantly granted to white customers. As the interviewee explained, “[t]hose who managed to get their fees waived were typically white baby boomers, but to my memory, the people who got their fees waived were mainly white. The Black customers wouldn’t typically ask for their fees to be waived.”¹⁷⁰

Additionally, some of the interviewees recounted making exceptions for wealthy clients to avoid losing their business, illustrating how perceived buying power and SES can influence decisions about accommodating special requests. For example, one clerk from a high-end shoe store highlighted the

¹⁶⁷ Interview with Interviewee no. 4 (Oct. 1, 2018) (transcript on file with *Northwestern University Law Review*).

¹⁶⁸ *Id.*

¹⁶⁹ Interview with Interviewee no. 6 (Oct. 7, 2018) (transcript on file with *Northwestern University Law Review*).

¹⁷⁰ Interview with Interviewee no. 3, *supra* note 164.

importance of maintaining relationships with high-spending customers, explaining that “in a high-end luxury, you didn’t want to lose a good client. If you have a good client come back in, more so with the women than men, saying, ‘You know, I’ve worn these shoes for a couple of weeks . . . but they’re just not really working for me,’ you’re not gonna lose someone spending \$25,000 a year on your store over a \$500 pair,” even if the return policy says otherwise.¹⁷¹ Another clerk noted that higher SES customers typically felt more entitled to accommodations and had “better manners” than customers at stores such as Target. According to this salesclerk, because such “[p]eople were more ‘up-scale,’” they “didn’t want to be made [a] fool of themselves.”¹⁷²

Research suggests that these types of perceptions can intersect with racial and gender biases, even if such biases are not explicitly acknowledged by the clerks. Indeed, evidence shows that Black consumers are typically perceived as less valuable customers in consumer markets,¹⁷³ with race and class being intricately linked in societal perceptions.¹⁷⁴ Clerks might (consciously or unconsciously) equate white customers with higher SES, thereby treating them more favorably under the assumption that they are more valuable to the store.¹⁷⁵

To conclude, the findings from the interviews illuminate the multifaceted decision-making process behind return-policy deviations, revealing a complex interplay of customer service, incentive structures, and discretionary power. Although clerks emphasize a commitment to customer satisfaction and store reputation, their accounts also hint at the subtle yet significant role of demographic perceptions in shaping their decisions. Despite explicit denials of bias, the patterns observed suggest that assumptions about race, gender, class, and expected customer behavior can

¹⁷¹ Interview with Interviewee no. 5 (Oct. 11, 2018) (transcript on file with *Northwestern University Law Review*).

¹⁷² Interview with Interviewee no. 9, *supra* note 162.

¹⁷³ See, e.g., Pittman, *supra* note 36, at 3, 4, 15–16.

¹⁷⁴ See, e.g., Ridgeway & Kricheli-Katz, *supra* note 129, at 304 (presenting how a “variety of social groups clustered together . . . found that ‘whites’ and ‘middle-class people’ were closely clustered, as were ‘Blacks’ and ‘blue-collar people’”); Penner & Saperstein, *supra* note 152, at 320 (finding that changes in how people were perceived racially were inextricably linked to and influenced by social-class position, and that people who experienced “decreases in status” in a given wave of a longitudinal survey were more likely to be classified by interviewers as Black rather than white).

¹⁷⁵ See, e.g., Freeman et al., *supra* note 133, at 2, 8 (finding that wearing clothing representative of a lower SES increases the likelihood that a person would be classified as Black rather than white); Brewster, *supra* note 153, at 3, 21–23; Brewster et al., *supra* note 153, at 506, 524 (finding servers are likely to discriminate against and withhold effort from Black patrons due to negative perception of the tipping practice of Black diners).

influence the fairness and consistency of return policy applications. This insight underscores the need for further exploration into how implicit biases may permeate retail practices, potentially contributing to the racial disparities observed in consumer treatment.

The Article now pivots to discussing the normative implications of these disparities, highlighting the need for policy interventions that address and mitigate the impacts of such discriminatory practices in the retail sector and beyond.

V. POLICY RESPONSES

The findings in this Article suggest that a key issue in consumer-contract discrimination lies in a salesclerk's ability to deviate from contractual terms, potentially influenced by implicit biases. One proposed solution is to prohibit any deviations from standard contracts or policies to prevent discrimination.¹⁷⁶ Admittedly, that idea, at least on its face, has merit. By prohibiting *all* discretionary deviations from form contracts, regulators could prevent sellers from discriminating against minority consumers. Prohibiting sellers from allowing their employees to deviate from policies or standard contracts, however, has several key disadvantages.

First, implementing and enforcing such a prohibition could be difficult. Sellers often have a vested interest in allowing employee discretion to maintain customer satisfaction and loyalty, which could easily lead to resistance against strict adherence to a new rule.¹⁷⁷ Moreover, enforcement agencies may not always know of these deviations, complicating their ability

¹⁷⁶ A similar proposal was raised in the context of taxicab tipping. The idea was floated following a study that revealed passengers systematically gave lower tips to minority drivers. In response, the researchers who conducted the study proposed a *service compris* ("tip included") regulation, whereby cab commissions would increase the metered price and require drivers to prominently display "tip included" decals. The researchers posited that prohibiting tipping in this manner would reduce passenger discrimination against minority drivers as well as driver discrimination against minority passengers. See Ian Ayres, Fredrick E. Vars & Nasser Zakariya, *To Insure Prejudice: Racial Disparities in Taxicab Tipping*, 114 YALE L.J. 1613, 1658–60 (2005).

¹⁷⁷ For further detail regarding the connection between employee discretion and customer satisfaction and loyalty, see, for example, Jean-Charles Chebat & Paul Kollias, *The Impact of Empowerment on Customer Contact Employees' Roles in Service Organizations*, 3 J. SERV. RSCH. 66 (2000); Avantika Tomar & Amit Dhiman, *Customer Oriented Bureaucracy & Flexibility: Interactions & Implications in the Services Industry Context*, 48 INDIAN J. INDUS. RELS. 43 (2012); and Samina Quratulain, Moh'd Ahmad Al-Hawari & Shaker Bani-Melhem, *Perceived Organizational Customer Orientation and Frontline Employees' Innovative Behaviors: Exploring the Role of Empowerment and Supervisory Fairness*, 24 EUR. J. INNOVATION MGMT. 533 (2021).

to investigate and act against noncompliant sellers. Ensuring compliance could thus present considerable challenges.¹⁷⁸

The second concern regarding an outright ban on deviations is that sellers might well respond by either decreasing quality or increasing prices.¹⁷⁹ To illustrate this point, consider how this might play out in the context of retail returns. Sellers could, for example, opt to strictly enforce a “no refund without a receipt” policy for everyone, or they could choose to offer a more accommodating “full refund even without a receipt” policy but raise the price of the item. The rub lies in the fact that sellers would have to offer the same prices and policy terms to *all* consumers and, depending on what the sellers choose, such regulation could disproportionately burden lower income consumers with higher prices or inferior terms (even if applied universally).¹⁸⁰

The final and perhaps most significant concern is that a strict prohibition on deviations might lead to unfair outcomes, where sellers enforce one-sided terms against all consumers, even when fairness or other circumstances call for accommodation. In other words, while discretion might leave room for bias, it also provides the flexibility often necessary to

¹⁷⁸ Enforcement concerns might be mitigated by allowing class actions and imposing punitive damages, which would enable collective consumer action for broader accountability and deter sellers through heightened sanctions. See John C. Coffee Jr., *Understanding the Plaintiff's Attorney: The Implications of Economic Theory for Private Enforcement of Law Through Class and Derivative Actions*, 86 COLUM. L. REV. 669, 679 (1986); Bryant Garth, Ilene H. Nagel & S. Jay Plager, *The Institution of the Private Attorney General: Perspectives from an Empirical Study of Class Action Litigation*, 61 S. CAL. L. REV. 353, 379, 383 (1988); Mitchell Polinsky & Steven Shavell, *Punitive Damages: An Economic Analysis*, 111 HARV. L. REV. 869, 874, 887–97 (1998). Regulators could also employ audit studies to detect and address impermissible policy deviations, similar to methods used in housing and mortgage-lending discrimination cases. See generally Lincoln Quillian, John J. Lee & Brandon Honoré, *Racial Discrimination in the U.S. Housing and Mortgage Lending Markets: A Quantitative Review of Trends, 1976–2016*, 12 RACE & SOC. PROBS. 13 (2020) (conducting an audit study to reveal housing discrimination). However, these measures, while potentially effective, are not flawless and may not guarantee full deterrence. See, e.g., Maria Glover, *The Structural Role of Private Enforcement Mechanism in Public Law*, 53 WM. & MARY L. REV. 1137, 1141–46 (2011) (noting that limitations on private enforcement mechanisms, such as class actions, may impair deterrence).

¹⁷⁹ See, e.g., Johnston, *supra* note 55, at 879–80, 884.

¹⁸⁰ Traditionally, the dilemma has been framed as a choice between raising prices or lowering quality. Yet, experiences from non-retail sectors suggest a third path: differentiated terms at varied prices, a strategy seen in the airline and mobile phone industries where consumers choose from options that balance flexibility and cost. See, e.g., Florence Derrick, *Travel on Your Terms with Flexible Flights*, SKYSCANNER (Mar. 12, 2023), www.skyscanner.net/news/flexible-travel-guide [https://perma.cc/4KQU-27VR]; Goitom Tesfom, Nancy J. Birch & Jeffrey N. Culver, *Switching Behavior of U.S. Mobile Phone Service Customers After Providers Shift from Contract to No Contract Mobile Phone Service Plans*, 33 J. RETAILING & CONSUMER SERVS. 154, 154 (2016). However, this approach requires sellers to effectively segment consumers by their preferences and willingness to pay, a practice not yet broadly adopted in most consumer markets.

reach just decisions.¹⁸¹ This conundrum harks back to the old “rules versus standards” debate,¹⁸² highlighting the inevitable trade-off between equitable outcomes and uniformity of application.¹⁸³ Indeed, it is doubtful whether a rule that would work in all situations could be devised.¹⁸⁴ Consider, for instance, a situation where a residential tenant misses rent payment due to a car accident resulting in hospitalization. Should the representative of the residential company be barred from waiving the late fee stipulated in the lease agreement under such circumstances? Likewise, imagine a hotel guest who asks for a late checkout due to a family emergency. Should the front desk clerk be prohibited from accommodating this request if the room is not reserved for the subsequent night? As consumer–seller interactions evolve to be more relational and extended, the importance of allowing for “discretionary justice” in contractual implementation becomes increasingly important.¹⁸⁵

Given the substantial challenges of strictly prohibiting deviations from standard contracts and policies, I briefly sketch three cumulative solutions: first, explicitly prohibiting discrimination in the execution of consumer contracts;¹⁸⁶ second, expanding public accommodations laws; and third, encouraging sellers to implement measures designed to reduce bias in contractual performance.¹⁸⁷

¹⁸¹ Cf. KENNETH C. DAVIS, *DISCRETIONARY JUSTICE: A PRELIMINARY INQUIRY* 20 (1970) (arguing that discretion is often desirable in the governmental context because “no governmental authority . . . can make a new policy or new law without exercising discretion”).

¹⁸² See, e.g., Kaplow, *supra* note 73, at 562–63 (explaining that rules involve “advance determinations of the law’s content,” whereas standards require “later determinations of the law’s content” and are therefore more costly for legislators and enforcement authorities); Kathleen M. Sullivan, *The Justices of Rules and Standards*, 106 HARV. L. REV. 22, 62, 66 (1992); Kennedy, *supra* note 73, at 1701.

¹⁸³ See, e.g., Sunstein, *supra* note 73, at 974–77.

¹⁸⁴ A certain amount of discretion is necessary in any decision-making system, including the legal system. See generally Kenneth Culp Davis, *Discretionary Justice*, 23 J. LEGAL EDUC. 56, 58 (1970) (“[D]iscretion, even unguided discretion, is an absolute necessity for every legal system.”); Antonin Scalia, *The Rule of Law as a Law of Rules*, 56 U. CHI. L. REV. 1175, 1185 (1989) (“[T]he establishment of broadly applicable general principles is an essential component of the judicial process . . .”).

¹⁸⁵ The literature on relational contracts is vast. The canonical works include, *inter alia*, Ian R. Macneil, *Relational Contract: What We Do and Do Not Know*, 3 WIS. L. REV. 483 (1985); Melvin A. Eisenberg, *Why There Is No Law of Relational Contracts*, 94 NW. U. L. REV. 805 (2000); and Richard E. Speidel, *The Characteristics and Challenges of Relational Contracts*, 94 NW. U. L. REV. 823 (2000).

¹⁸⁶ For a detailed discussion of the proposal to ban discrimination in the performance of consumer contracts, including its advantages and limitations, see Furth-Matzkin, *supra* note 22.

¹⁸⁷ For a detailed discussion of the bias-reduction strategies sellers can implement to reduce bias in contractual performance, see Furth-Matzkin, *supra* note 23, at 1185–97.

A. Prohibiting Discrimination in Contractual Performance

Prohibiting discrimination in the performance of sellers' contracts or policies could be achieved by broadening the interpretation of 42 U.S.C. § 1981, which prohibits racial discrimination in making and enforcing contracts, to cover discrimination during the performance of contracts, not just during contract formation.¹⁸⁸ Section 1981 ensures equal rights "to make and enforce contracts,"¹⁸⁹ and a 1991 amendment clarified that this protection extends to all aspects of contractual relationships, including performance.¹⁹⁰ Yet, despite what appears to be unequivocal language, courts have traditionally limited its scope to contract formation.¹⁹¹

Recent case law suggests a potential judicial shift toward recognizing discrimination during the execution of contracts.¹⁹² Courts should take this opportunity to explicitly broaden the interpretation of § 1981 to include protections against discriminatory practices in the performance phase of consumer contracts.¹⁹³ Such an expansion would ensure more robust and comprehensive safeguards, aligning the law with its goal of preventing discrimination.¹⁹⁴

B. Expanding Public Accommodation Laws

The federal public accommodations statute contains a list of public places, including restaurants, theaters, and hotels, in which discrimination is prohibited, but it fails to mention retail stores.¹⁹⁵ Unfortunately, courts have so far interpreted that list as exhaustive, refusing to include retail stores under its scope.¹⁹⁶ Meanwhile, the state-level response to this issue presents a

¹⁸⁸ See 42 U.S.C. § 1981(a) (stipulating that all persons shall have equal rights to "make and enforce contracts . . . as is enjoyed by white citizens").

¹⁸⁹ *Id.*

¹⁹⁰ Civil Rights Act of 1991, Pub. L. No. 102-166, § 101, 105 Stat. 1071, 1072 (codified at 42 U.S.C. § 1981(b)) (clarifying that these protections extend to the "making, performance, modification, and termination of contracts, and the enjoyment of all benefits, privileges, terms, and conditions of the contractual relationship").

¹⁹¹ See Furth-Matzkin, *supra* note 22, at 17–18.

¹⁹² *Id.* at 19–20.

¹⁹³ *Id.*

¹⁹⁴ For a more thorough examination of this proposal, see *id.* at 19–43.

¹⁹⁵ See 42 U.S.C. § 2000a.

¹⁹⁶ See, e.g., *Newman v. Piggie Park Enters., Inc.*, 377 F.2d 433, 436 (4th Cir. 1967) ("Retail stores, food markets, and the like were excluded from [Title II of the 1964 Civil Rights Act] for the policy reason that there was little, if any, discrimination in the operation of them." (citing 110 CONG. REC. 6533 (1954) (statement of Sen. Humphrey))), *aff'd*, 390 U.S. 400 (1968). For a critique of this approach, see Joseph William Singer, *No Right to Exclude: Public Accommodations and Private Property*, 90 NW. U. L. REV. 1283, 1288–89 (1996).

fragmented picture. For example, in the retail context, forty-four states have laws explicitly prohibiting racial discrimination in retail spaces, but six states—Alabama, Florida, Georgia, Mississippi, North Carolina, and Texas—lack such explicit protections, with Mississippi allowing retail stores to refuse service at their discretion.¹⁹⁷ In the financial markets context, some states have recently begun enacting (or are considering enacting) legislation requiring financial institutions to provide customers with “fair access” to financial services. Tennessee and Florida, for example, have enacted fair access laws that came into effect in July 1, 2024, and July 1, 2023, respectively, and which apply to national and state banks and insurers.¹⁹⁸ At least eight other states—Arizona, Georgia, Idaho, Indiana, Iowa, Kentucky, Louisiana, and South Dakota—are also considering fair access bills, some of which would apply to payment processors, payment networks, and credit card companies and networks in addition to banks and insurers.¹⁹⁹

These inconsistencies highlight a significant challenge: ensuring uniform and comprehensive civil rights protections across the United States. Furthermore, even if all states adopted these laws, it remains unclear whether courts will interpret them as prohibiting discrimination in contractual performance. A straightforward solution is thus to amend the federal public accommodations statute to explicitly include retail stores. By doing so, we can create a more equitable framework that protects all consumers, regardless of race, and reinforces the commitment to civil rights across the nation.

C. Encouraging Sellers to Adopt Bias-Reducing Measures

Given the role that implicit biases and stereotypes most likely play in shaping contractual implementation in consumer markets, sellers should be

¹⁹⁷ Hila Keren, *Separating Church and Market: The Duty to Secure Market Citizenship for All*, 12 U.C. IRVINE L. REV. 911 (2022); *Public Accommodations Laws: 50-State Survey*, JUSTIA (last reviewed Mar. 2024), <https://www.justia.com/civil-rights/public-accommodations-laws-50-state-survey/> [https://perma.cc/B8Z3-KTY4]; Singer, *supra* note 196, at 1435.

¹⁹⁸ See H.B. 2100, 113th Gen. Assemb., Reg. Sess. (Tenn. 2023) (codification in Tennessee); H.B. 3, 2023 Leg., Reg. Sess. (Fla. 2023) (codification in Florida); see also *HB 3 Implementation for Financial Services Providers*, FLA. OFF. OF FIN. REGUL., <https://lofr.gov/sitepages/documents/HB3ImplementationforFinancialServicesProviders.pdf> [https://perma.cc/V34L-QQ2F] (providing an introduction to the Florida bill).

¹⁹⁹ See S.B. 1167, 56th Leg., 2d Reg. Sess. (Ariz. 2024); H.B. 1205, 157th Gen. Assemb., 2023–2024 Reg. Sess. (Ga. 2024); H.B. 669, 67th Leg., 2d Reg. Sess. (Idaho 2024); S.B. 28, 123d Gen. Assemb., 2d Reg. Sess. (Ind. 2024); H.B. 2409, 90th Gen. Assemb., 1st Sess. (Iowa 2024); H.B. 452, 2024 Leg., Reg. Sess. (Ky. 2024); H.B. 914, 2024 Leg., Reg. Sess. (La. 2024); H.B. 1247, 99th Leg. Assemb., Reg. Sess. (S.D. 2024).

incentivized (or required) to adopt various measures known to reduce racial biases in decision-making.²⁰⁰

Four key measures in particular warrant careful consideration. First, to help employees recognize and reduce their implicit biases, sellers could implement debiasing (or diversity) training programs.²⁰¹ The effectiveness of such training, particularly over the long term, depends on tailoring content to specific needs, providing ample time for comprehensive sessions, and using tools such as the IAT to raise awareness of personal biases.²⁰² Second, holding employees accountable for their decisions can play a crucial role in mitigating bias.²⁰³ This can be accomplished through monitoring sales staff, requiring them to provide justifications for discretionary choices, and establishing objective criteria to guide their decision-making processes.²⁰⁴

²⁰⁰ See Furth-Matzkin, *supra* note 23, at 1185–97 (proposing such measures).

²⁰¹ Diversity and debiasing trainings draw from social psychological research on implicit bias. They underscore the premise that individuals can address and rectify their biases once they are made aware of their existence and are motivated to take corrective measures. See, e.g., Richardson & Goff, *supra* note 150, at 120 (“Reducing [implicit biases] is more likely when individuals are aware of the potential for biased decisionmaking and are aware of the possibility of safeguarding against the influence of implicit bias.”); Susan T. Fiske, *Stereotyping, Prejudice, and Discrimination*, in 2 THE HANDBOOK OF SOCIAL PSYCHOLOGY 357, 357 (Daniel Gilbert, Susan Fiske & Gardner Lindzey eds., 4th ed. 1998) (examining how conscious awareness of implicit biases, combined with motivation to change, can lead individuals to develop control strategies that reduce discriminatory behaviors); Anthony G. Greenwald, Nilanjana Dasgupta, John F. Dovidio, Jerry Kang, Corinne A. Moss-Racusin & Bethany A. Teachman, *Implicit-Bias Remedies: Treating Discriminatory Bias as a Public-Health Problem*, 23 PSYCH. SCI. PUB. INT. 7, 11–12 (2022); Elizabeth Levy Paluck, Roni Porat, Chelsey S. Clark & Donald P. Green, *Prejudice Reduction: Progress and Challenges*, 72 ANN. REV. PSYCH. 533 (2021). Prominent examples of diversity training in the consumer sector include Starbucks’s nationwide closure for diversity training in 2018 following an incident involving the police and Black patrons, and Sephora’s sensitivity workshop in 2019, prompted by accusations of racial profiling against Black consumers. See, e.g., Shep Hyken, *Starbucks Closes 8,000 Stores for Racial Bias Training—Is It Enough?*, FORBES (June 1, 2018, 4:05 AM), <https://www.forbes.com/sites/shephyken/2018/06/01/starbucks-closes-8000-stores-for-racial-bias-training-is-it-enough/?sh=1150be572831> [<https://perma.cc/H8UM-576S>]; Jonah E. Bromwich, *Sephora Will Shut Down for an Hour of Diversity Training Tomorrow*, N.Y. TIMES (June 4, 2019), <https://www.nytimes.com/2019/06/04/style/sephora-will-shut-down-for-an-hour-of-diversity-training-tomorrow.html> [<https://perma.cc/H6XV-3SKC>].

²⁰² See, e.g., Alexandra Kalev, *The Impact of Diversity Training Programs in the Workplace and Alternative Bias Reduction Mechanisms*, OXFORD RSCH. ENCYC. BUS. & MGMT., Sept. 20, 2023, at 1 (suggesting ways for workplaces to make diversity training more effective); Patricia G. Devine & Tory L. Ash, *Diversity Training Goals, Limitations, and Promise: A Review of the Multidisciplinary Literature*, 73 ANN. REV. PSYCH. 403, 421 (2022).

²⁰³ See, e.g., Richardson & Goff, *supra* note 150, at 119 (observing that accountability can reduce implicit bias “because people exercise more care when they know their decisions are monitored and will have to be explained, and because thinking more carefully and deliberately helps to debias”).

²⁰⁴ See, e.g., Eric L. Uhlmann & Geoffrey L. Cohen, *Constructed Criteria: Redefining Merit to Justify Discrimination*, 16 PSYCH. SCI. 474, 479 (2005) (finding that using objective criteria to guide decision-making can reduce bias); Calvin K. Lai, Kelly M. Hoffman & Brian A. Nosek, *Reducing Implicit*

Given that this study's findings suggest a tentative link between discretionary language in return policies and heightened racial bias, implementing measures to limit such discretion appears promising for reducing bias in customer interactions.

Third, increasing diversity among sales personnel can reduce implicit biases and improve interactions with consumers from historically marginalized groups.²⁰⁵ Finally, artificial intelligence (AI) can be a powerful tool for reducing bias in decision-making.²⁰⁶ Stores could, for instance, utilize AI to monitor employees and enhance accountability.²⁰⁷ AI's ability to detect patterns of bias has already proven effective in areas such as judicial and medical decision-making.²⁰⁸ In the context of contractual performance,

Prejudice, 7 SOC. & PERSONALITY PSYCH. COMPASS 315, 322 (2013) (finding that using objective criteria to guide decision-making can reduce bias); Monica Biernat & Melvin Manis, *Shifting Standards and Stereotype-Based Judgments*, 66 J. PERSONALITY & SOC. PSYCH. 5, 19 (1994) (finding that objective criteria reduce stereotypical thinking).

²⁰⁵ See, e.g., Jerry Kang & Mahzarin R. Banaji, *Fair Measures: A Behavioral Realist Revision of "Affirmative Action,"* 94 CALIF. L. REV. 1063, 1109 (2006) (surveying the literature showing that diversity reduces bias and proposing that employers hire "debiasing agents"—women and people of color—to reduce bias in the workplace); Jamillah B. Williams, *Accountability as a Debiasing Strategy: Testing the Effect of Racial Diversity in Employment Committees*, 103 IOWA L. REV. 1593, 1601 (2018). It is important to note, however, that recent Supreme Court rulings that disallowed race-conscious decisions in university admissions could potentially discourage sellers from adopting diversity initiatives. See Lauren Aratani, *Where Now for Workplace Diversity After Court's Affirmative Action Ruling?*, GUARDIAN (July 2, 2023), <https://www.theguardian.com/law/2023/jul/02/workplace-diversity-dei-supreme-court-affirmative-action> [<https://perma.cc/652T-4VHP>] (noting the concern that companies will "second-guess their DEI initiatives after the court's decision"). But cf. Kenji Yoshino & David Glasgow, *What SCOTUS's Affirmative Action Decision Means for Corporate DEI*, HARV. BUS. REV. (July 12, 2023), <https://hbr.org/2023/07/what-scotuss-affirmative-action-decision-means-for-corporate-dei> [<https://perma.cc/5R3A-LFDJ>] (noting that, despite the Supreme Court's recent ruling, companies may still legally advance DEI initiatives in the workplace).

²⁰⁶ AI can potentially also play a role in diminishing bias in discretionary decisions. Indeed, AI is rapidly advancing and has already proven effective in detecting patterns of bias in areas such as judicial and medical decision-making. See, e.g., S. Mills, S. Cota & C.R. Sunstein, *AI, Behavioral Science, and Consumer Welfare*, 46 J. CONSUMER POL'Y 387, 392–93 (2023) (surveying AI's expanding and fast-growing capabilities in detecting biases); Jon Kleinberg, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig & Sendhil Mullainathan, *Human Decisions and Machine Predictions*, 133 Q.J. ECON. 237, 288–89 (2018).

²⁰⁷ On the expanding and fast-growing capabilities of AI in detecting biases more generally, see, for example, Mills et al., *supra* note 206; and also see Kleinberg et al., *supra* note 206, at 289.

²⁰⁸ Kleinberg et al., *supra* note 206, at 288–89; see Carlo Reverberi, Tommaso Rigon, Aldo Solari, Cesare Hassan, Paolo Cherubini, GI Genius CADx Study Group & Andrea Cherubini, *Experimental Evidence of Effective Human–AI Collaboration in Medical Decision-Making*, SCI. REPS., Sept. 2, 2022, at 1, 5–8 (finding that a combination of AI and human-endoscopist input resulted in more accurate diagnoses than that of either input individually); see also Anto Čartolovni, Ana Tomičić & Elvira Lazić Mosler, *Ethical, Legal, and Social Considerations of AI-Based Medical Decision-Support Tools: A Scoping Review*, INT'L J. MED. INFORMATICS, Mar. 14, 2022, at 1, 4 (surveying literature on AI-assisted medical diagnoses and finding that "[s]ome studies emphasize . . . the implications of AI-based decision support tools in improving diagnostic and prognostic efficiency").

retailers could harness AI's pattern-detection capabilities to identify discriminatory deviations from contractual terms or policies, thus reducing the human bias that often influences discretionary decisions.

Admittedly, AI should be used with caution in this regard for several reasons. First, AI often relies on historical data, which may itself contain inherent biases.²⁰⁹ For instance, if algorithms are fed data from past discriminatory practices (such as return refusals), they might inadvertently perpetuate these biases.²¹⁰ Second, correlations in the data—such as those between income, education, and race—can obscure underlying racial biases.²¹¹ Third, the inclusion of human judgment in the decision-making loop can reintroduce biases.²¹² Yet, despite these challenges, AI, when applied carefully and critically, has the potential to play an important role in reducing discrimination in contractual performance through monitoring and accountability.²¹³

Policymakers could incentivize sellers to adopt those anti-bias strategies by offering legal protections or imposing liability on those who fail to take action against known biases.²¹⁴

CONCLUSION

This Article underscores the persistence of racial discrimination in the performance of consumer contracts, particularly in retail settings. The findings from the audit study reveal that Black customers continue to face significant disadvantages in post-contractual interactions, such as returning goods, compared to their white counterparts. As suggested by the clerk interviews, this disparate treatment is likely rooted in implicit racial biases

²⁰⁹ The literature on algorithmic bias is vast. For prominent examples, see Aziz Z. Huq, *Racial Equity in Algorithmic Criminal Justice*, 68 DUKE L.J. 1043, 1051–52, 1079 (2019), which discusses possible implications of biased training data for criminal justice tools using AI; Pauline T. Kim, *Data-Driven Discrimination at Work*, 58 WM. & MARY L. REV. 857, 860–61 (2017); Andrew D. Selbst, *Disparate Impact in Big Data Policing*, 52 GA. L. REV. 109, 120 (2017); and Tal Z. Zarsky, *Understanding Discrimination in the Scored Society*, 89 WASH. L. REV. 1375, 1389 (2014).

²¹⁰ Talia Gillis & Jann L. Spiess, *Big Data and Discrimination*, 86 U. CHI. L. REV. 459, 487 (2019); Sandra G. Mayson, *Bias In, Bias Out*, 128 YALE L.J. 2218, 2224 (2018); Kleinberg et al., *supra* note 206, at 239–40.

²¹¹ See, e.g., Talia Gillis, *The Input Fallacy*, 106 MINN. L. REV. 1175, 1223 (2022).

²¹² Talia Gillis, Bryce McLaughlin & Jann Spiess, *On the Fairness of Machine-Assisted Human Decisions 2* (Sept. 2023) (unpublished manuscript), <https://wwws.law.northwestern.edu/research-faculty/events/colloquium/law-economics/documents/on-the-fairness-of-machine-assisted-human-decisions.pdf> [<https://perma.cc/D79H-ECNX>].

²¹³ Certain paths forward have already been proposed. See, e.g., Crystal S. Yang & Will Dobbie, *Equal Protection Under Algorithms: A New Statistical and Legal Framework*, 119 MICH. L. REV. 291, 376 (2020) (discussing how to promote fairness in machine-assisted human decision-making processes).

²¹⁴ See Furth-Matzkin, *supra* note 23, at 1197–99.

and entrenched stereotypes about racial differences in buying power, expectations, and behavior. To more robustly combat retail race discrimination in contractual performance, policymakers could consider broadening the reach of existing legislation to explicitly outlaw racially discriminatory practices in the execution of seller contracts, as well as encouraging sellers to adopt strategies that have been empirically shown to reduce bias.

ONLINE APPENDIX

The Online Appendix is hosted on Digital Commons and can be accessed at <https://scholarlycommons.law.northwestern.edu/nulr/vol119/iss5/1/>.

