Explaining Racial Disparities in Personal Bankruptcy Outcomes*

Bronson Argyle† Sasha Indarte‡ Benjamin Iverson§ Christopher Palmer¶

March 2024

Abstract

We document substantial racial disparities in consumer bankruptcy outcomes and investigate the role of racial bias in contributing to these disparities. Using data on the near universe of US bankruptcy cases and a deep-learning imputed measures of race, we show that Black filers are 16 and 3 percentage points (pp) more likely to have their bankruptcy cases dismissed without any debt relief in Chapters 13 and 7, respectively. We uncover strong evidence of racial homophily in Chapter 13: Black filers are 7 pp more likely to be dismissed when randomly assigned to a White bankruptcy trustee. To interpret our findings, we develop a general decision model and new identification results relating homophily to bias. Our homophily approach is particularly useful in settings where traditional outcomes tests for bias are not feasible because the decision-maker's objective is not well defined or the decision-relevant outcome is unobserved. Using this framework and our homophily estimate, we conclude that at least 37% of the 16 pp dismissal gap is due to either taste-based or inaccurate statistical racial discrimination.

Keywords: personal bankruptcy, racial disparities, racial bias, homophily

*First Version: January 11, 2021. Code implementing our race imputation procedure is available here. We thank Libby Cahoon, Sam Earnest, Noah Gillespie, Brice Green, Tanvi Jindal, Declan Kahley, Eric Myzelev, SeSe Nguyen, Damon Petersen, Sarah Raviola, Eleni Sivers, Luca Sivers, Simon Sivers, and Paige Stevenson for their excellent research assistance. We thank discussants Ellora Derenoncourt, Paul Goldsmith-Pinkham, Arpit Gupta, Brittany Lewis, Feng Liu, Erik Mayer, and Michaela Pagel. For helpful feedback, we also thank Ivan Canay, Will Dobbie, Hanming Fang, Peter Hull, Ben Keys, Paul Mohnen, Ed Morrison, Matt Notowidigdo, and Will Dobbie. We also thank conference and seminar participants at the BFI Conference on Discrimination in the 21st Century, CFPB, Federal Reserve Bank of Philadelphia, Colorado Finance Summit, QSIDE Colloquium, Chicago Booth Data and Welfare in Household Finance, SITE Financial Regulation, Wharton, Federal Reserve Bank of San Francisco, University of Nebraska, Opportunity Insights, UCSB LAEF Conference of Racial Inequality, University of Oregon, NBER Household Finance, Housing, Household Debt, and the Macroeconomy Conference (Chicago), CEPR Household Finance, ASU. We acknowledge financial support from the Alfred P. Sloan Foundation through an NBER Household Finance Small Grant and The Wharton School's Jacobs Levy Equity Management Center for Quantitative Financial Research. Palmer acknowledges funding from NSF CAREER Grant 1944138. This project received exempt status from MIT's IRB (E-5130).

†Brigham Young University; bsa@byu.edu
‡University of Pennsylvania; aindarte@wharton.upenn.edu
§Brigham Young University; ben_iverson@byu.edu
¶Massachusetts Institute of Technology and NBER; cjpalmer@mit.edu

1 Introduction

Each year, close to one million US households enter consumer bankruptcy, receiving debt relief worth more than the resources given through all state unemployment insurance programs combined (Lefgren, McIntyre and Miller, 2010).[1] Given its scale, a first-order policy concern is understanding whether and why the bankruptcy system works less well for different households. For example, Lefgren and McIntyre (2009) show that bankruptcy rates in zip codes that are predominantly Black are nearly twice as large as White zip codes, and Kiel and Fresques (2017) find that personal bankruptcy filers from Black zip codes are more than twice as likely to have their cases dismissed (without any debt relief) than observationally similar filers from White zip codes.

Racial disparities in financial outcomes are large and widespread. For example, the median wealth of White households is more than ten times that of Black and Hispanic households ($171,000 versus $17,000, SCF, 2016). Minorities also pay higher interest rates than Whites with similar observable characteristics (Ghent, Hernandez-Murillo and Owyang, 2014; Bayer, Ferreira and Ross, 2018; Butler, Mayer and Weston, 2023). Racial disparities in consumer bankruptcy, a key part of the social safety net, may further compound existing economic and financial disparities by limiting access to this major source of debt relief.

This paper presents new facts on racial disparities in consumer bankruptcy and provides the first evidence of the role of racial bias in contributing to these disparities. Our analysis uses a new dataset containing the universe of US bankruptcy cases over the past two decades, containing detailed data on tens of millions of bankruptcy cases. To investigate the role of bias, we develop a decision model and new identification results that formalize when and how homophily between bankruptcy filers and their legal counterparts (judges and trustees) can both signal the presence of racial bias and quantify the share of observed disparities due to racial bias.[2] Although the role of bankruptcy trustees is under-explored in both the economics and law literatures, we find bankruptcy trustees play an important role in determining case outcomes. We focus on bias on the part of bankruptcy trustees (legal official involved in the bankruptcy process).

[1] Bankruptcy-system leniency is positively related to debtor wages, credit access, homeownership, and longevity (Dobbie and Song, 2015; Dobbie, Goldsmith-Pinkham, and Yang, 2017; Dobbie, Mahoney, Goldsmith-Pinkham, and Song, 2020).

[2] As we discuss below, homophily in our context is when there are differences in the treatment of filers when the filer and the legal decision-maker are of the same versus different races.

Understanding what drives disparities across groups of filers, particularly with regard to race, is important for assessing the efficacy of bankruptcy policy and has important implications for other economic outcomes. Bankruptcy is a frequently used form of social insurance—over 10% of US households have filed for bankruptcy at least once (Stavins, 2000; Keys, 2018). If in practice the institution of bankruptcy works poorly for certain disadvantaged groups, it could exacerbate wealth and welfare gaps caused by racial bias in financial and labor markets. For example, (Ganong, Jones, Noel, Greig, Farrell and Wheat, 2020) show that Black households cut consumption 50 percent more than White households in response to a similar-sized income shock. If Black households on average lack access to liquidity to smooth consumption—potentially due to fewer labor market opportunities or less access to savings and credit—the insurance provided through the bankruptcy system could be particularly valuable for these households. However, if racial bias diminishes the benefits that these groups receive in bankruptcy, the system is potentially providing less relief to those individuals that need it most. If the primary mechanism for individual debt relief in the US exhibits racial bias, such biases could amplify the effects of other racial disparities in financial markets and ultimately have important differential effects on wealth and wellbeing across groups.

To test for racial bias in the personal bankruptcy system, we assemble a nationwide dataset of detailed bankruptcy outcomes where meaningful demographic characteristics of the judge, trustee, and filer can be either observed or more confidently imputed than in previous work. Our primary analysis subsample is the near universe of consumer bankruptcy cases since 2008, containing over 13 million bankruptcy cases. We impute race by training a deep learning model extending Kotova (2021) on voter registration data from seven states that contain names, addresses, and self-reported race for over 36 million individuals. We then test for and analyze the role of bias in explaining disparities in several bankruptcy outcomes, many of which have not been studied before because of data limitations. Traditional bankruptcy outcomes include dismissal (when the judge rejects the filing), chapter selection (Chapter 13 is less generous), bankruptcy refiling rates, and discharged debts. Richer outcomes not usually available to bankruptcy researchers include conversion (when the judge forces a conversion from Chapter 7 to Chapter 13), net debt forgiveness (defined as discharged assets minus payments out of seizable assets), home valuation, which debts are discharged, and whether the court made filing eligibility exceptions to time-since-filing, income means-testing thresholds, or asset holding thresholds.

Our first finding is that Black filers are 16 percentage points more likely to have their personal bankruptcy petitions dismissed in court without any debt relief in Chapter 13—see Figure 1 for overall dismissal rates by race. This rate is 27% of the average Chapter 13 dismissal rate. In Chapter 7, Black filer's cases are dismissed 3 percentage points more often, which is 101% of the Chapter 7 dismissal rate among White filers. Conditional on a rich array of fixed effects and case-level controls, these gaps drops to 9 and 0.5 percentage points (for Chapters 13 and 7, respectively).

Next, we examine how the racial disparity in dismissal rate varies with trustee race (homophily). For Chapter 13, we find that when Black filers are randomly assigned to a White trustee, their dismissal rate rises by 7 percentage points. When assigned to a non-White trustee, the Chapter 13 dismissal rate gap falls to one percentage point (and is statistically insignificant). For Chapter 7, we estimate a precise null effect of trustee race on the racial disparity in dismissal rates.

To guide the interpretation of our homophily results, we develop a general decision model and econometric results relating homophily to bias. We formalize conditions that are sufficient and necessary for homophily to identify the relative difference in racial bias across DMs of different races, yielding a test for the presence of bias. The generality of our framework makes it readily usable to study bias and homophily in other contexts, such as lending, real estate transactions, bail, jury convictions, or hiring. A key advantage of our approach is that it does not require that the econometrician observes (or even knows) the outcomes over which potentially biased decision-makers (DMs) optimize. Notably, this means that our homophily test may be used in situations where a Becker-style outcome test (e.g., Becker, 1957, 1993; Knowles, Persico and Todd, 2001; Arnold, Dobbie and Yang, 2018; Canay, Mogstad and Mountjoy, 2020; Hull, 2021) is infeasible due to either a lack of data on outcomes or due to decisions being abstract (i.e., difficult to define) or complex (e.g., depending on multiple outcomes that may interact). To further quantify the direction and size of racial bias requires additional assumptions. To this end, we present a menu of increasingly strong assumptions that lead to increasingly informative lower bounds on the share of disparities due to racial bias.

We employ our homophily approach to detect bias because, unlike bail decisions where judges have a clear objective to minimize pre-trial misconduct, bankruptcy decisions are more complex and subjective. When determining whether to dismiss a bankruptcy case, judges and trustees optimize over multiple competing objectives, some of which are hard-to-measure and highly subjective. For example, these include maximizing creditor recoveries, denying improper creditor claims, minimizing bankruptcy fraud (e.g., are misreported assets due to intentional fraud or procedural error?), correctly measuring filer assets and income, and (in Chapter 13) confirming a feasible repayment plan, and whether filer hardship is "beyond their control."[3] Judge and trustee objectives may also be abstract if they derive utility from, for example, a sense of correctly applying the law when making subjective evaluations.

In the model, a decision-maker (DM) makes a binary choice that affects another individual (e.g., a trustee chooses whether to dismiss a filer's case). She maximizes her expected utility when making this choice, solving a prediction problem where she forecasts how her decision will impact a vector of outcomes that she cares about (detecting fraud, giving access to debt relief, her own compensation, etc.). Her decision can be influenced by three forms of racial bias. First, taste-based racial bias can arise if the race of the filer alters the utility she receives when a particular outcome is realized (e.g., a trustee dislikes fraud more when committed by a Black filer). Second, inaccurate statistical discrimination can lead to dismissals if DMs have inaccurate beliefs about how a filer's race predicts the outcomes they care about. Third, accurate statistical discrimination can influences dismissals if filer race (accurately) predicts differential likelihoods of outcomes DMs value. Despite our use of the terms accurate or inaccurate statistical discrimination, we stress that any statistical discrimination is potentially problematic, especially given that the non-race characteristics used in statistical discrimination models are themselves often the product of historical discrimination (Spriggs, 2020). We refer to the effect of taste-based and inaccurate statistical discrimination as "$\beta\mu$-racial bias" and the combined effect of all three biases as "total racial bias."

We show that homophily identifies the relative differences in DMs' total racial bias if, and only

[3]In bankruptcy cases, judges and trustees evaluate the accuracy and completeness of a petitioner's reported assets, liabilities, income sources, and expenses—and whether the petitioner has strategically manipulated any of these variables. Perceived manipulation can warrant modifying the calculations using these variables that determine the amount petitioners must pay their creditors to successfully discharge their debt, making it harder to avoid dismissal due to missed payments. Perceptions of egregious manipulation can trigger immediate dismissal without debt relief. Additionally, Chapter 13 filers that encounter financial hardships during their (three to five year) repayment plan can request a hardship discharge, which requires a subjective evaluation by a court over the extent to which the petitioner's hardship was "out of their control" and makes repayment infeasible. These abstract criteria are not as readily measurable as pre-trial misconduct.

if, dismissal decisions exhibit "parallel disparities."[4] In the bankruptcy context, parallel disparities means that if (counterfactually) Black filers were instead White but still had their same non-race characteristics, the disparity in Black-White dismissals would be the same for filers assigned to either Black or White trustees. We then show that if, and only if, decisions exhibit parallel accurate statistical discrimination, then the relative differences in total racial bias equal the difference in $\beta\mu$-racial bias across DMs. This second assumption means that, in the absence of $\beta\mu$-racial bias, the effect of filer race on dismissal decisions would be the same for filers assigned to either Black or White trustees. We propose two tests to check for failure of these assumptions. The first is a standard balance test, which checks whether DM race predicts filer race or non-race characteristics. The second tests whether non-race characteristics *differentially* predict dismissals across cases assigned to White versus non-White DMs. The latter could detect if, for example, White DMs react differently to non-race characteristics that are potentially correlated with race. Applying these tests in our setting suggests that violation of these key assumptions is unlikely.

These theoretical results make homophily a potentially a powerful tool for isolating $\beta\mu$-racial bias from total racial bias. This framework leads to two implications of the substantial homophily we find in Chapter 13. First, it implies that trustees of at least one racial type exhibit some form of $\beta\mu$-racial bias (i.e., taste-based or inaccurate statistical discrimination). Second, if non-White trustees are either unbiased or biased against Black filers on average, then at least 37% of the 16 percentage point dismissal disparity is due to $\beta\mu$-racial bias.

Related Literature In this section, we briefly contextualize our findings in related literatures on personal bankruptcy, racial bias in credit markets and institutions, and law and economics.

First, we build on a growing literature on racial disparities in household financial outcomes through a new focus on disparities and bias in personal bankruptcy. Prior work documents disparities in lending outcomes, such as minorities experiencing lower loan approval rates, higher interest rates, and higher rates of CFPB complaints (Munnell, Tootell, Browne and McEneaney, 1996; Reid, Bocian, Li and Quercia, 2017; Bayer, Ferreira and Ross, 2018). More recently, a growing literature documents the challenges faced by algorithmic advances in underwriting struggle

[4]We term this assumption parallel disparities because of its similarity to the parallel trends assumption required for identification using a difference-in-differences estimator.

to eliminate racial disparities and bias in credit outcomes (Bartlett, Morse, Stanton and Wallace, 2019; Fuster, Goldsmith-Pinkham, Ramadorai and Walther, Forthcoming; Morse and Pence, 2020; Blattner and Nelson, 2020).

Studying bankruptcy in particular, Braucher, Cohen and Lawless (2012) find that Black households file for Chapter 13 at a much higher rate.[5] Their experimental evidence suggests that attorney steering plays a role in explaining this disparity, with attorneys are more likely to recommend that clients with Black-sounding names file under Chapter 13 than otherwise identical filers with White-sounding names. In this paper, we contribute new evidence on disparities across a range of bankruptcy outcomes, document the role of trustees in shaping these disparities, and quantify the role of racial bias. A key difference in this paper relative to previous work is that we focus on documenting disparities *after* and individual has entered bankruptcy while holding constant all filer characteristics that existed at the time of the bankruptcy filing. As such, the racial differences that we find arise mostly due to the bankruptcy system itself rather than choices that consumers make prior to filing.

Second, we build on a law and economics literature exploring the importance of decision-maker characteristics in legal outcomes. In the context of bankruptcy, court congestion and inexperienced judges lead to worse bankruptcy outcomes, such as lower creditor recovery rates (Iverson, 2018; Iverson, Madsen, Wang and Xu, 2020). Other work on bias in the legal system finds evidence of racial bias in bail decisions (Arnold, Dobbie and Yang, 2018; Arnold, Dobbie and Hull, 2020). Additionally, juror race, gender, and political ideology impact conviction rates (Anwar, Bayer and Hjalmarsson, 2012, 2019a,b). We contribute by highlighting the role of trustee bias in shaping disparities in personal bankruptcy outcomes. Bankruptcy trustees have received little attention in prior work, and our findings suggest that they have a significant influence on bankruptcy outcomes, similar to that of judges. Recently, Morrison, Pang and Zytnick (2019) finds that, in several cities, lawyers appear to help clients strategically time their bankruptcy filings to improve their chances of obtaining a lenient trustee, suggesting that lawyers believe that trustees can dramatically impact bankruptcy outcomes. Importantly, such steering is unlikely a source of

[5]Chapter 13 is generally considered a "worse" form of bankruptcy for several reasons. First, Chapter 13 can require filers to make larger repayments to creditors (statute requires it be no less in Chapter 13 than what creditors would receive in Chapter 7). Second, Chapter 13 filers are less likely to receive a discharge at the conclusion of their case. Third, the Chapter 13 discharge is not received until completion of the payment plan, which is most often five years after filing. See Section 2 for background on the personal bankruptcy system.

bias in our empirical analysis. We document that trustee race does not predict filer race nor a variety of case and non-race characteristics. We are also able to account for trustee differences in overall leniency by including trustee fixed effects.

Third, this paper makes a methodological contribution to the literature on detecting and quantifying racial bias. Our homophily approach may be used to complement a Becker-style outcome test (e.g., Becker, 1957, 1993; Knowles, Persico and Todd, 2001; Arnold, Dobbie and Yang, 2018; Canay, Mogstad and Mountjoy, 2020; Hull, 2021) or in situations when such a test is infeasible due to unobserved or abstract DM objectives. Homophily between agents and DMs has been widely-studied in many contexts, such as police stops and searches (Anwar and Fang, 2006), jury convictions (Anwar, Bayer and Hjalmarsson, 2012), and mortgage lending (Jiang, Lee and Liu, 2021; Frame, Huang, Mayer and Sunderam, 2022). However, there is limited prior econometric and theoretical work guiding the interpretation of homophily. A notable exception is Anwar and Fang (2006), which presents results in which homophily can test for bias. Our framework generalizes their decision model to allow for complex decisions and for DMs to make prediction errors, which can result in inaccurate statistical discrimination. We also formalize identification results *quantifying* the impact of bias.

There are several limitations to our framework. We show that homophily may fail to detect bias when it is present, but homophily reliably indicates bias when it is detected.[6] In this sense, homophily is a conservative test for the presence of bias. And in general, our results quantifying bias partially identify the influence of bias, rather than point identify it. However, it is possible to obtain a sharp lower bound, as we do in our application. Additionally, although our homophily framework does not strictly require random assignment of decision-makers, in practice, this is likely necessary for the identifying assumptions to be credible.

The rest of the paper proceeds as follows. Section 2 presents relevant institutional background on personal bankruptcy in the US. Section 3 presents our decision model and econometric results.We describe our data and present descriptive facts on personal bankruptcy outcomes by groups in Section 4. Section 5 presents our results, and Section 6 concludes.

[6]Formally, our test is inconsistent (underpowered) but has exact size.

2 Background: Personal Bankruptcy in the US

2.1 The Costs and Benefits of Personal Bankruptcy

Nearly one million households every year seek to discharge consumer debts by filing for Chapter 7 or Chapter 13 Bankruptcy. Bankruptcy can help households cope with financial distress—for example, stemming from job loss or medical expenses—by reducing required debt payments and preventing wage garnishment. In doing so, bankruptcy offers households an implicit form of insurance that can help them better smooth consumption across states of the world (for evidence on the insurance value of bankruptcy see Livshits, MacGee and Tertilt, 2007; Chatterjee, Corbae, Nakajima and Ríos-Rull, 2007; Indarte, 2023; Dávila, 2020)(for evidence on the insurance value of bankruptcy see Livshits et al., 2007; Chatterjee et al., 2007; Indarte, 2023; Dávila, 2020). The scale of the debt relief offered under Chapters 7 and 13 is substantial, totaling $187 billion in a typical year.[7] During the Great Recession, the annual debt write-downs provided by bankruptcy were similar in size to the annual transfers from unemployment insurance and larger than those of measures like the Home Affordable Modification Program (Auclert, Dobbie and Goldsmith-Pinkham, 2019).

Receiving a debt discharge through bankruptcy can benefit filers in many dimensions. Financially, filers typically see better credit scores and credit access in the years after filing compared to insolvent non-filers (Albanesi and Nosal, 2018). Filers that receive a discharge (versus those whose case is dismissed) also experience higher earnings, lower foreclosure rates, higher homeownership rates, and lower mortality rates (Dobbie and Song, 2015; Dobbie, Goldsmith-Pinkham and Yang, 2017). Consistent with smoothing and stabilizing consumption, Auclert, Dobbie and Goldsmith-Pinkham (2019) find that access to bankruptcy increased employment by nearly 2% during the Great Recession.

Even in situations where the filing is dismissed and debt is not ultimately discharged, there can still be beneficial effects of filing. Filing a voluntary petition for bankruptcy with the court triggers an automatic stay of creditor's legal ability to pursue outstanding debts (both secured and unsecured obligations), which allows filers to keep assets such as a vehicle or a home. This stay can be especially meaningful in minimizing the disruption of financial distress.

Filing for bankruptcy also entails a number of costs. Court, attorney, and mandatory debt

[7] Source: Annual BAPCPA report (Tables 1A and 1D).

counseling fees average $1,400 in Chapter 7 and $3,400 in Chapter 13 (GAO, 2008). Additionally, filers can be required to make payments to creditors out of assets (Chapter 7 bankruptcy) or out of disposable income (Chapter 13 bankruptcy). Non-monetary costs like stigma may also be an important deterrent to filing (Indarte, 2023). In the long-term, the "bankruptcy flag" that appears on a filer's credit report for seven to ten years can depress credit access (Musto, 2004; Dobbie et al., 2020; Herkenhoff et al., 2019; Gross et al., 2020). Filing today also costs filers the option to file in the near future, as discharges can only be granted every two to eight years.[8] If a filer's petition is dismissed, not only do they not receive the debt discharge, but they will still bear many of the costs of bankruptcy—including receiving a "bankruptcy flag" on their credit report.

2.2 The Bankruptcy Process

Below we describe the bankruptcy process, highlighting the role played by trustees and judges as well as the relevant differences between Chapter 7 and Chapter 13 bankruptcy. To initiate bankruptcy proceedings, a filer/petitioner first must complete schedules thoroughly detailing their assets, liabilities, income sources, and expenses. Within 15 days of submitting this paperwork, the file must also provide proof of completing a credit counseling course. The course helps filers prepare a budget and explore options for repaying their debts. The course also offers an assessment of the feasibility of repaying debt, which judges can take into account when ruling in bankruptcy cases.

After completing these two steps, filers then participate in a Meeting of Creditors (341 Hearing). This meeting is run by the bankruptcy trustee. If the filer fails to attend their case may be dismissed; if the filer has hired a lawyer, they will also attend. Creditors may attend but rarely do so (Elias and Bayer, 2017). This meeting is an important step for the trustee to form a recommendation to the judge and detect fraud.

The trustee compares the paperwork detailing the filer's financial data to financial documents (tax returns, bank statements, auto titles, etc.) to ensure its accuracy and to detect fraud. The trustee must verify that the filer qualifies to file under the Chapter that they have chosen (to file for Chapter 7, the filer's income must be below the state's median). Additionally, the trustee

[8] Chapter 7 filers must wait eight years to file again under Chapter 7 and four years to file under Chapter 13. Chapter 13 filers must wait two years to file again under Chapter 13 and six years to file under Chapter 7.

may question the filer about the reasonability of asset valuations and expenses, the ability of the filer to sustain a high enough income to afford monthly payments under a proposed Chapter 13 repayment plan, and whether misreported values reflect fraud or innocent mistakes. In Chapter 7 bankruptcy, the trustee gains the power to sell the filer's assets with value in excess of state-specific exemption limits.[9] Other forms of fraud the trustee will look for are transfers of assets that were intended to reduce the value of nonexempt assets and credit-financed purchases where the filer had no intention of repaying the debt. Within 60 days after the Meeting of Creditors, the filer must complete a debtor education course, which emphasizes budgeting and rebuilding credit after bankruptcy.

Bankruptcy cases terminate in one of three ways: discharge, conversion, or dismissal. If the filer succeeds in receiving a discharge at the conclusion of their case, their debts are wiped out after making required payments to creditors. The main differences between Chapters 7 and 13 are the timing and amounts of payments to creditors. In Chapter 7, the filer pays the value of assets in excess of their state's exemption limits. This occurs soon after the Meeting of Creditors if the trustee and creditors have no objections.

In Chapter 13, the filer attends a court hearing, with both the judge and trustee present, to confirm their proposed repayment plan. Statute requires that the sum of Chapter 13 payments are at least as high as what the creditor would have received under Chapter 7 (the value of non-exempt assets). The payments can be higher, in which case they equal the filer's disposable income (income minus "necessary" expenses). Chapter 13 filers make monthly payments for three to five years, and the discharge is not received until after the completion of payments.

Chapter 13 filers may also receive an early discharge of their debt if they encounter financial hardship that makes their initial payment plan infeasible. When the filer experiences a major income loss or rise in expenses – for example, due to the closing of a plant or illness – the judge and trustee may determine that the filer qualifies for a hardship discharge. If the hardship is not seen as beyond the control of the filer or insufficiently severe, missed payments may instead result in a dismissal (Elias and Bayer, 2017). However, the judge and trustee may also approve of a modified Chapter 13 plan.

[9]The filer can purchase nonexempt assets in order to retain them. For example, if the filer's home equity exceeded its exemption limit by $15,000 but their retirement savings were fully exempt, they could use their retirement savings to pay the $15,000 to keep their home.

When cases end in conversion, the filer is forced to file under a different Chapter. A conversion from Chapter 7 to 13 typically happens if the filer's income is above the state's median, which usually disqualifies them from Chapter 7. It may also happen if the judge and trustee believe that the filer can feasibly repay more of their debt under Chapter 13. Conversions from Chapter 13 to 7 typically occur when the judge and trustee believe the proposed Chapter 13 repayment plan is infeasible. After a case is converted, the filer still has the chance to successfully obtain a debt discharge in a new case under a different Chapter.

An unsuccessful outcome for the filer is a dismissal, in which case the filer does not receive a debt discharge, though they have benefited from the automatic stay against creditors mentioned above. If the case is dismissed without prejudice they can refile again immediately. If dismissed with prejudice, the filer typically has to wait one year to file again, but the exact timing is at the discretion of the judge (Elias and Bayer, 2017), and creditors can also bring suit granting exception of the automatic stay if the subsequent filing is ruled by the judge to be "in bad faith." Cases are commonly dismissed for several reasons: fraud, failure to complete mandatory educational classes, failure to file forms or submit documents, failure to pay court fees, missing the Meeting of Creditors, perceived infeasibility of the Chapter 13 payment plan, and missed Chapter 13 payments. When filers simply make a procedural mistake, they are more likely to receive a dismissal without prejudice.

Scope for Bias. Bankruptcy trustees and judges face several subjective evaluations. If racial bias can affect their perceptions of honesty and hardship, trustees may suggest – and judges may opt – to dismiss cases at different frequencies for otherwise similar filers of different races. Race may be made especially salient to trustees, who meet face-to-face with filers (of all chapters) during the Meeting of Creditors and any court hearings. In Chapter 7, filers rarely need to attend a court hearing with the judge. But race may be more salient to a judge in Chapter 13 cases, which require a confirmation hearing to approve the Chapter 13 plan.

The trustee plays a central role in evaluating whether a filer's actions constitute *intentional* fraud as opposed to a *procedural error*. This includes assessing whether a transfer of property was intended to reduce the value of nonexempt assets, the filer intended to repay a recent credit-financed purchase, misreported income was an oversight or a mistake, or an event merits a hardship discharge. Additionally, the trustee may disagree over the reasonableness in counting some expenses as necessary and forecasts for future income. If bias leads to more expensive Chapter 13 plans for Black filers, they may be more likely to have their case dismissed due to perceived infeasibility or actual difficulties in making payments. Trustees make recommendations for discharges and dismissals based on their evaluations, but ultimately a judge must form their own opinions to decide the outcome of a case.

3 Decision Model and Econometric Framework

This section presents the framework that guides the interpretation of our empirical analysis. We ground our exposition in the context of personal bankruptcy. However, we emphasize that the framework developed here can be applied to study bias in a variety of other settings. Notably, our framework can be used in settings where bias is otherwise difficult to study with outcome tests (e.g., the tests formulated in Becker, 1957; Arnold et al., 2018; Hull, 2021; Bohren et al., 2022) due to the complexity and/or lack of observability of the outcomes that the decision-maker optimizes over.

We begin by developing a model of the bankruptcy dismissal decision. In the context of the model, we define three sources of discrimination that may influence the decision: accurate statistical discrimination, inaccurate statistical discrimination, and taste-based discrimination.[10] We then define causal parameters and discuss their interpretation as measures of bias. With these definitions in hand, we show how differences in racial disparities across legal decision-makers with different races (homophily) can be used to test for the presence of racial bias and partially identify the share of observed disparities attributable to racial bias.

3.1 Decision Model

3.1.1 Notation and Setup

A bankruptcy case is characterized by a random set

[10] Our focus on these three sources of bias parallels that of Hull (2021), who relates similar notions of these forms of bias to outcome tests.

$$(J, R_J, I, R_I, X, D, Y_1, Y_0)$$

where J indexes an individual decision-maker (DM) who has race $R_J \in \{b, w\}$, in the context of bankruptcy, this could be either a judge or a trustee. The other party in the case is a filer, indexed by I with race $R_I \in \{b, w\}$ and non-race characteristic $X \in \mathbb{R}$. We simplify this setting by having only two races: Black (b) and white (w) and by having only a scalar non-race characteristic. Our results are not sensitive to this choice and can be extended to accommodate a larger, finite number of racial identities as well vectors of non-race characteristics. Additionally, without loss of generality, we could alternatively interpret x as a noisily-measured signal of a non-race characteristic (but for simplicity, below we do not explicitly model measurement error in x).

The DM selects a binary decision $D \in \{0, 1\}$; in our context $D = 1$ denotes dismissing the bankruptcy case and $D = 0$ denotes not dismissing the case. The DM's decision influences the "outcome" of the case $Y_D = Y_0 + (Y_1 - Y_0)D$. We allow this outcome to be a vector containing multiple "sub-outcomes", that is, Y_D is an $m \times 1$ vector with $m \geq 1$. In the context of bankruptcy, sub-outcomes could include whether the filer receives debt relief, whether the filer committed fraud in the bankruptcy process, whether the filer makes the required bankruptcy payments, or the compensation received by the DM. We use lower case letters (j, r_j, i, r_i, x) to refer to specific parties and their characteristics. We refer to DMs with she/her/hers pronouns and filers with he/him/his pronouns.

3.1.2 The Decision-Maker's Problem

The DM's utility depends on the outcome of the case. She chooses whether to dismiss in order to maximize her expected utility. Formally, she solves

$$\max_{d \in \{0,1\}} E_j[u(Y_d; j, r_i, x) | r_i, x].$$

Her utility function $u(\cdot)$ is parameterized by (j, r_i, x). This flexibly allows utility functions $u(\cdot)$ to vary with the DM's identity j. This means that, for example, DMs can have different preferences for preventing bankruptcy fraud. It also allows utility to vary with the filer's race r_i and non-race characteristic x for a given outcome and DM. This would allow, for example, for a given DM to prefer when low-income or White filers receive debt relief (relative to high-income or Black filers).

The expectations operator E_j denotes an expectation calculated using DM j's beliefs about the likelihood that various possible outcomes Y_d are realized. Ex ante, Y_d is unknown to the DM. And our setup also does not require that the DM ever fully learns Y_d after making her decision. This means that she could be concerned with reducing bankruptcy fraud, but fraud could be committed without her confirming that it did occur. We also do not require that the DM has correct beliefs (i.e., her beliefs do not have to coincide with the true, objective probabilities). We assume that she observes the filer's race and non-race characteristic and may condition her expectation on these variables. She may therefore (correctly or incorrectly) believe that, for example, the likelihood of the filer completing all plan payments is predicted by the filer's race and employment status. We do not explicitly model other non-race characteristics that are *unobserved* by the DM and influence outcomes. Rather, we capture this in a simplified way by allowing the case outcome Y_d to be uncertain.

It will be convenient to focus on the difference in utility under the dismissal choices facing the DM:

$$\Delta(j, r_i, x) \equiv u(Y_1; j, r_i, x) - u(Y_0; j, r_i, x).$$

With this notation in hand, we can write the DM's optimal decision as:

$$D(j, r_i, x) = \mathbf{1}\{E_j[\Delta(j, r_i, x)|r_i, x] \geq 0\}.$$

The DM dismisses the case when her expected utility from doing so is weakly higher.

3.1.3 The Sources of Bias and their Influence on Decisions

Bias Decomposition. We next highlight how different forms of bias can influence decisions. Note that we use the terms "bias" and "discrimination" interchangeably. Distinguishing types of bias is important because the nature of discrimination affects the efficacy of different policy tools that could be used to reduce bias. Defining different forms of bias in the context of the model also clarifies the economic meaning of the identification assumptions and results ahead.

We begin by decomposing the DM's payoff into three components that highlight channels through which distinct forms of bias can impact dismissal decisions. The first component corresponds to prediction error:

$$\mu(j, r_i, x) \equiv E[\Delta(j, r_i, x)|r_i, x] - E_j[\Delta(j, r_i, x)|r_i, x].$$

Above, μ equals the difference in the DM's expected payoffs under the true conditional distribution of (Y_1, Y_0) versus her believed distribution. Importantly, this difference holds constant the conditioning variables and parameterization of the DM's utility function. Differential prediction error by filer race can result in *inaccurate statistical discrimination* or "stereotyping" (e.g., as in Bordalo, Coffman, Gennaioli and Shleifer, 2016; Bohren, Haggag, Imas and Pope, 2020). Specifically, this form of discrimination can arise when prediction errors change when the race of the filer changes, holding constant the DM and the filer's non-race characteristic (i.e., $\mu(j, b, x) \neq \mu(j, w, x)$). Our definition of prediction error mirrors that of Canay, Mogstad and Mountjoy (2020), defined in the context of our generalized decision problem.

The second component corresponds to *taste-based discrimination*:

$$\beta(j, r_i, x) \equiv E[\Delta(j, w, x)|r_i, x] - E[\Delta(j, b, x)|r_i, x].$$

Here, β captures how the DM's expected payoff changes when her utility is evaluated for a white versus Black filer. The two differenced terms both use the same, correct conditional distributions when evaluating the expectations. They also both have same non-race characteristic x parameterizing the utility function. A nonzero β therefore arises when there exists some outcomes Y_d where the DM's payoff varies with the the race of the filer. For example, they dislike missed payments more when the filer is white.

The third component relates to *accurate statistical discrimination* for a reference group. Here, we use White filers as a reference group (this choice does not affect the results).

$$E[\Delta(j, w, x)|r_i, x].$$

The above expression is the true expected utility gain from dismissal (Δ) when utility is parameterized for a White filer. Crucially, the expectation above conditions on the filer's true race, r_i.

Accurate statistical discrimination can arise when changing whether we condition on the filer being White or Black changes the DM's true expected payoff, when utility is parameterized by the reference group (i.e., $E[\Delta(j,w,x)|b,x] \neq E[\Delta(j,w,x)|w,x]$). Accurate statistical discrimination occurs when race is correlated with factors that affect the DM's utility, leading the DM to partially base her optimal decision on the filer's race. For example, suppose that the DM's utility depends on preventing bankruptcy fraud and that the propensity to commit fraud is not directly observed by the DM but is correlated with the filer's race. In this case, the DM may choose to dismiss a case for a filer of a given race when she would not dismiss the case if the filer had a different race since the filer's race gives an indication the fraud is more likely to occur.

We decompose DM's expected payoff into these three components:

$$E_j[\Delta(j,r_i,x)|r_i,x] = E[\Delta(j,w,x)|r_i,x] - \mu(j,r_i,x) - \mathbf{1}[r_i = b]\beta(j,b,x).$$

A positive accurate statistical discrimination term (first component) implies that the filer's race and non-race characteristic (accurately) predict that the DM would prefer to dismiss when utility is parameterized by a white filer. A positive μ means that the DM's prediction errors lead her to underestimate the true utility gain from dismissing, decreasing her preference for dismissal. A positive β (taste-based discrimination) indicates that the DM has a lower expected utility gain when dismissing a Black filer, holding constant the other facts of the bankruptcy case, decreasing her preference for dismissing Black filers. Negative μ and β terms would instead increase the DM's preference for dismissal.

The above decomposition does not a priori rule out taste-based discrimination against White filers. The β term disappears in the equation above when the filer is White ($r_i = w$). However, we can rewrite the above instead using Black filers as the reference group for the accurate statistical discrimination term; this version would instead contain a $\mathbf{1}[r_i = w]\beta(j,w,x)$ term.

Defining Racial Bias. We present two definitions of bias, defined at the case-level.

Definition 1: Total racial bias.

(a) *A bankruptcy case is said to exhibit* total racial bias *if* $D(j,b,x) \neq D(j,w,x)$.

(b) *A bankruptcy case is said to exhibit* total racial bias against Black filers *if* $D(j,b,x) > D(j,w,x)$.

(c) *A bankruptcy case is said to exhibit* total racial bias favoring white filers *if* $D(j, b, x) < D(j, w, x)$.

The term "total" emphasizes that this definition allows any of the three forms of bias to influence the dismissal decision. Whether DM biases *alter* the DM's decision when the filer's race changes is central in this definition. Note that a DM may still have differing payoffs when dismissing two otherwise identical White and Black borrowers, but that need not alter their decision. Such a case, under our definition, would not be *exhibiting* total bias. This notion of total bias is similar to the "local bias" of Canay, Mogstad and Mountjoy (2020) in that allows for a given DM to be biased in some cases and not in others, or biased against Black or White filers depending on the non-race characteristic. Our definitions of bias are also similar to that of Hull (2021) in that that they relate to the *decision* made by the DM (as opposed to the outcomes over which the DM optimizes).

Our second definition of bias pertains to the influence of taste-based and inaccurate statistical-discrimination. To introduce it we decompose the DM's decision into two components:

$$D(j, r_i, x) = \widetilde{D}(j, r_i, x) + \widetilde{\beta\mu}(j, r_i, x)$$

where

$$\widetilde{D}(j, r_i, x) \equiv \mathbf{1}\{E[\Delta(j, w, x)|r_i, x] \geq 0\}$$

$$\widetilde{\beta\mu}(j, r_i, x) \equiv D(j, r_i, x) - \widetilde{D}(j, r_i, x).$$

The $\widetilde{D}$ term characterizes the decision that a DM would make if only influenced by accurate statistical discrimination with a White filer reference group. The $\widetilde{\beta\mu}$ term captures the net influence of both taste-based and inaccurate statistical discrimination.When $\widetilde{\beta\mu} = 1$, this indicates that the case was dismissed but would *not* have been dismissed in the absence of taste-based and inaccurate statistical discrimination. When $\widetilde{\beta\mu} = 0$, taste-based and inaccurate statistical discrimination on net did not alter the dismissal decision. Lastly, when $\widetilde{\beta\mu} = -1$, taste-based and inaccurate statistical discrimination resulted in a case avoiding dismissal that otherwise would have been dismissed.

We next define a second case-level definition of racial bias.

Definition 2: $\beta\mu$-racial bias.

(a) *A bankruptcy case is said to exhibit* $\beta\mu$-racial bias *if* $\widetilde{\beta\mu}(j,b,x) \neq \widetilde{\beta\mu}(j,w,x)$.

(b) *A bankruptcy case is said to exhibit* $\beta\mu$-racial bias against Black filers *if* $\widetilde{\beta\mu}(j,b,x) > \widetilde{\beta\mu}(j,w,x)$.

(c) *A bankruptcy case is said to exhibit* $\beta\mu$-racial bias favoring Black filers *if* $\widetilde{\beta\mu}(j,b,x) < \widetilde{\beta\mu}(j,w,x)$.

This second type of racial bias reflects the combined influence of taste-based and inaccurate statistical discrimination.

Policy Significance of Different Forms of Bias. Understanding the importance of different components of bias is important for identifying which policies can most effectively reduce racial bias. Black-white dismissal disparities due to accurate statistical discrimination can be alleviated through several types of policies. One variety targets reducing the predictiveness of race for outcomes Y_D that the DM optimizes over. For example, policies that make Black and White filer income risk more similar would make their likelihood of completing bankruptcy payments more similar. Another policy instead offers DMs financial incentives based on filer race to bring their decisions into alignment. A third type of policy that could reduce bias from accurate statistical discrimination is to standardize the DM process, limiting the role of subjective predictions. At one extreme, this could be implemented if DMs were blinded to filer race and based dismissal decisions on formulaic rules based only on non-race characteristics (in effect, automating decision-making). Such rules can still result in racial disparities due to correlations between race and non-race characteristics (Fuster, Goldsmith-Pinkham, Ramadorai and Walther, Forthcoming).

How can inaccurate statistical discrimination be reduced? The rules-based approach mentioned above could also help lessen the impact of this form of discrimination. Differences in DM ability to accurately predict filer outcomes could be erased if all DMs were compelled to use the same decision rule. However, the design of a such a rule could nonetheless still reflect prediction error among the parties designing the rule. Additional policy variants that could alleviate inaccurate statistical discrimination include providing DMs data or guidelines to help reduce prediction errors. If own-group prediction error also tends to be smaller, another policy type that could reduce this form of bias would be increasing diversity among DMs. A more controversial variant on this policy would be to explicitly match DMs and filers on the basis of race. Matching could also create unequal workloads if Black DMs are underrepresented and Black filers are over-represented. If DMs make more prediction errors when facing a large number of cases, this policy could potentially backfire.

Taste-based racial discrimination could also be alleviated by increasing DM diversity or matching DMs and filers based on race. Another policy tool that could reduce taste-based racial bias is debias training of DMs, which aims to reduce implicit bias. Sanctioning DMs exhibiting explicit bias can also limit the influence of taste-based racial discrimination. Ahead, we develop a set results isolating the influence of taste-based and inaccurate statistical discrimination. This focus is motivated by two reasons. The first is that the policy prescriptions vary with the nature of discrimination. The second is that it is possible to modify dismissal decisions to reduce taste-based and inaccurate statistical discrimination and at the same time *increase* average DM utility (net of preferences for discrimination). In contrast, modifying dismissal decisions to reduce accurate statistical discrimination generally reduces average DM utility.

3.1.4 Causal Parameters of Interest

We present two causal parameters that quantify the influence of racial bias on bankruptcy dismissal decisions. We first define the causal effect of filer race on Black filer's outcomes:

$$\delta^{ATT} \equiv E[D(j, b, x) - D(j, w, x) | r_i = b].$$

This estimand describes on average how Black filers' race changes whether their case is dismissed when holding constant their non-race characteristic x. A positive value means that Black filers on average experience more dismissals than they otherwise would if the only characteristics of theirs that changed was their race. The model highlights three channels through which such discrimination can arise. The above estimand corresponds to the total effect on Black filer's dismissals of all three channels. More specifically, it is the average amount of Black filers' cases altered by *total racial bias*, where total bias against Black filers enters positively and total bias favoring Black filers enters negatively.

Our second estimand focuses on the extent to which $\beta\mu$-racial bias (i.e., operating through both taste-based and inaccurate statistical discrimination) impacts Black filer dismissals:

$$\delta^{\beta\mu} \equiv E[\widetilde{\beta\mu}(j, b, x) - \widetilde{\beta\mu}(j, w, x) | r_i = b].$$

Recall that $\widetilde{\beta\mu}$ describes how taste-based and inaccurate statistical discrimination alters dismissal decisions. The estimand above therefore describes the average causal impact of the filer's race on their dismissal outcome via these two sources of bias.

3.2 Identification: Detecting and Quantifying Bias with Homophily

We now turn to how homophily can help researchers detect and quantify bias. Homophily is a widely-studied phenomenon in a variety of settings, often interpreted to be informative about bias.[11] The econometric framework we introduce below illustrates assumptions that make it possible to draw such conclusions from observational data.

In what follows, we assume that the researcher observes the filer and DM's races and the dismissal decisions (R_I, R_J, D). Notably, we do not require that the researcher observes the outcome vector over which the DM optimizes (Y_D), the filer's non-race characteristic (x), nor how DM preferences vary with their identity J ($u(\cdot)$). Indeed, our framework allows for a setting where the researcher does not know *what* outcomes the DM cares about (i.e., the components of the outcome vector Y_D). It is in this sense that our framework applies to settings where DMs make abstract or complex decisions. For simplicity, we abstract away from non-race characteristics observable to both the DM and researcher, as the framework below can be readily modified to condition on additional observables.

Identification Challenges. We highlight two distinct identification challenges. The first is that simply comparing differences in dismissal rates across filer outcomes does not identify the average total racial bias experienced by Black filers (δ^{ATT}):

$$E[D|r_i = b] - E[D|r_i = w] = \underbrace{E[D(j, b, x) - D(j, w, x)|r_i = b]}_{=\delta^{ATT}} + \underbrace{E[D(j, w, x)|r_i = b] - E[D(j, w, x)|r_i = w]}_{= \text{selection bias}}.$$

[11] For example: police stops/searches (Anwar and Fang, 2006), jury convictions (Anwar, Bayer and Hjalmarsson, 2012), and mortgage lending (Jiang, Lee and Liu, 2021; Frame, Huang, Mayer and Sunderam, 2022).

The selection bias term can be nonzero if non-race filer characteristics, observed by the DM but not observed by the researcher, are correlated with filer race and influence the DM's decision. For example, if Black filers face greater risk of job loss, DMs that value avoiding filers failing to make plan payments may be more likely to dismiss Black filers not (directly) because of their race but because of employment risk.

The second identification challenge is separating $\beta\mu$-racial bias from total racial bias. Total racial bias is the sum of accurate statistical racial discrimination and $\beta\mu$-racial bias:

$$\delta^{ATT} = E[\widetilde{D}(j,b,x) - \widetilde{D}(j,w,x)|r_i = b] + \delta^{\beta\mu}.$$

Even with an estimator for δ^{ATT}, we cannot identify $\delta^{\beta\mu}$ without either estimates of accurate statistical discrimination or the assumption that it equals zero. Estimating accurate statistical discrimination directly would be especially difficult as it describes the *hypothetical* decisions that DMs would make in the absence of taste-based and inaccurate statistical discrimination.

3.2.1 Homophily and Identifying Assumptions

We next show how homophily can help researchers learn about either total racial bias (δ^{ATT}) or $\beta\mu$-racial bias ($\delta^{\beta\mu}$). We start by making minimal assumptions and then show how increasingly strong assumptions can help researchers obtain sharper conclusions about the presence and size of racial bias. By gradually adding assumptions, we aim to present a "menu" of assumptions that researchers could use to determine what conclusions they can draw from homophily estimates. We also suggest several tests that can help falsify some of the assumptions.

The Homophily Estimand. Let τ denote the homophily estimand:

$$\tau \equiv \left\{E[D|r_i = b, r_j = w] - E[D|r_i = w, r_j = w]\right\} - \left\{E[D|r_i = b, r_j = b] - E[D|r_i = w, r_j = b]\right\}.$$

To minimize notation, going forward we will write the conditional expectation $E[D|r_i, r_j] = E_{r_i r_j}[D]$. That is, $E_{bw}[D]$ denotes the average dismissal rate conditional on having a Black filer and White

DM. Under this notation, the homophily estimand is

$$\tau = \{E_{bw}[D] - E_{ww}[D]\} - \{E_{bb}[D] - E_{wb}[D]\}.$$

The homophily estimand describes the differences in racial disparities among cases assigned to White versus Black DMs.

Identifying Total Racial Bias Under Parallel Disparities. Our first results highlight a necessary and sufficient condition for the homophily estimand to identify the average *difference* in racial bias (towards Black filers) among White versus Black DMs. Let

$$\delta_B^{ATT} \equiv E_{bb}[D(j,b,x) - D(j,w,x)]$$

$$\delta_W^{ATT} \equiv E_{bw}[D(j,b,x) - D(j,w,x)],$$

which correspond to average total bias experienced among Black filers assigned to Black and White DMs (respectively).

The key assumption for our first identification result is bankruptcy cases exhibit "parallel disparities", defined below.

Assumption 1: Parallel Disparities.

$$E_{bw}[D(w)] - E_{ww}[D(w)] = E_{bb}[D(w)] - E_{wb}[D(w)].$$

To minimize notation, we suppress the dependence of the dismissal decision on DM identity j and the filer's non-race characteristic x. The parallel disparities assumption is so-named because it resembles the "parallel trends" assumption from the difference-in-difference estimator. Indeed, it may be helpful to note that the homophily estimand τ is analogous to a difference-in-differences estimator where the first difference is across filer races and the second difference is across DM races. As opposed to the parallel trends assumption requiring that counterfactual time trends to be equal across Black and White DMs, parallel disparities requires counterfactual racial disparities to be equal across Black and White DMs. Intuitively, parallel disparities says that the difference in Black and White filer outcomes due to non-race characteristics, which may be correlated with race, is the same among filers assigned to either White or Black DMs.

Parallel disparities is a weaker assumption than assuming random assignment of DMs to cases. However, random assignment of DMs—that is, assignment independent of DM and filer characteristics (x, j)—implies parallel disparities. Unlike random assignment, parallel disparities would allow, for example, Black trustees to be more likely to face Black filers or White trustees more likely to face unemployed filers. However, it would need to be the case that these correlations result only in level differences in outcomes and not differences in *disparities* across Black and White filers. It may be difficult to be certain that such correlations are not also resulting in differential disparities. Thus, in practice researchers would likely prefer to have the stronger assumption of random DM assignment met.

Parallel disparities allows for level differences in outcomes. It does not impose that the non-race characteristics influencing the dismissal decision (x) are the same across Black and White filers, allowing for different average dismissal rates across Black and White filers. It also allows for DM leniency to be correlated with DM race. For example, Black DMs could be more lenient on average, dismissing fewer cases. There are two main scenarios that could violate parallel disparities.

The first arises if non-race characteristics x are differently correlated with filer race across filers facing Black versus White DMs. We would most likely worry about this scenario if filers could choose their DMs or vice versa. For example, suppose older White filers prefer (and could choose) to work with White DMs, while age doesn't change Black filer preferences. If DMs are generally more lenient toward older filers, parallel disparities could fail to hold. To test for violations of this sort, we recommend researchers conduct balance tests to verify that filer race and non-race characteristics do not predict DM race.

The second scenario occurs if Black and White DM decisions respond differently to non-race characteristics that are correlated with race. For example, if Black DMs are more lenient to low-income filers, and this characteristic is correlated with filer race, parallel disparities could fail to hold. This violation would lead to the homophily coefficient estimating a combination of racial bias and bias towards other characteristics that are correlated with race. The equivalent challenge in a standard difference-in-differences context would be if two events occurred simultaneously in the treated group, preventing separately identifying the the effect of either event. To test for this kind of violation, we recommend researchers interact non-race characteristics with DM race to look for evidence of systematic differences in decision-making. For our analysis, we present the results of these tests in Section 5.3. We show that filer race and non-race characteristics do not predict DM race and that White and non-White DMs do not react significantly differently to non-race characteristics.

Proposition 1: Identification of Difference in Average Total Racial Bias. *If parallel disparities (Assumption 1) holds, the homophily estimand identifies the average difference in total racial bias between Black and White DMs. That is,*

$$\tau = \delta_W^{ATT} - \delta_B^{ATT}.$$

Proof. First, rewrite the estimand in terms of potential outcomes:

$$\tau = \left\{E_{bw}[D(b)] - E_{ww}[D(w)]\right\} - \left\{E_{bb}[D(b)] - E_{wb}[D(w)]\right\}.$$

Note that we can add and subtract additional potential outcome terms to rewrite the two terms in brackets, respectively, as

$$\underbrace{E_{bw}[D(b) - D(w)]}_{=\delta_W^{ATT}} + E_{bw}[D(w)] - E_{ww}[D(w)]$$

$$\underbrace{E_{bb}[D(b) - D(w)]}_{=\delta_B^{ATT}} + E_{bb}[D(w)] - E_{wb}[D(w)].$$

With the above, we can apply the parallel disparities assumption to rewrite the homophily estimand as simply:

$$\tau = \delta_W^{ATT} - \delta_B^{ATT}. \quad \square$$

Intuitively, under parallel disparities, homophily overcomes the initial selection problem by differencing out the impact of non-race characteristics on decisions. With only the parallel disparities assumption, the homophily estimand could reflect either accurate statistical, inaccurate statistical, or taste-based discrimination. We next introduce a second assumption, "parallel accurate statistical discrimination" that further affects the interpretation of the homophily estimand.

Assumption 2: Parallel Accurate Statistical Discrimination (PASD).

$$E_{bw}[\widetilde{D}(b) - \widetilde{D}(w)] = E_{bb}[\widetilde{D}(b) - \widetilde{D}(w)]$$

Assumption 2 (PASD) states that, on average, the effect of changing Black filers' race on the dismissal decision that would arise in the absence of prediction error and taste-based discrimination is the same across Black and White DMs. In other words, if Black and White DMs made decisions based purely on accurate statistical discrimination, the effect of the filer's race on dismissal would be similar across both groups of DMs. Recalling the definition of $\widetilde{D}(j, r_i, x) = \mathbf{1}\{E[\Delta(j, w, x)|r_i, x]\}$ highlights that varying the filer's race only changes the (true) conditional distribution used by the DM to predict the likelihood of various outcomes in the PASD assumption.

Possible violations of this assumption are similar in nature to those that we would worry about regarding parallel disparities. For example, suppose that low-income status is correlated with filer race and Black low-income filers are more likely to have a Black DM (but not low-income white filers). PASD could fail to hold in this case if low-income status and race (correctly) jointly predict different case outcomes Y_D. Random assignment of DMs (independent of filer characteristics) would similarly alleviate this concern. Hence, the balance test suggested for attempting to falsify parallel disparities can also provide evidence against violations of PASD.

The second scenario again relates to differences in DM preferences for outcomes by race. If DMs only made decisions based on accurate statistical discrimination, PASD could fail to hold if Black and White DMs have different values of outcomes that are (accurately) predicted by filer race. Under a null hypothesis of no taste-based or inaccurate statistical racial discrimination (i.e., no $\beta\mu$-racial discrimination), testing for nonzero interactions between DM race and filer non-race characteristics in predicting dismissals could also falsify PASD. Under the alternative hypothesis (nonzero $\beta\mu$-racial discrimination among some DMs), the same test would only fail to falsify PASD in the knife-edge case where $\beta\mu$-racial bias perfectly offsets differences in DM preferences for dismissal under accurate statistical discrimination.

The additional assumption of PASD results in the homophily estimand identifying the average difference in $\beta\mu$-racial discrimination between Black and white DMs. We formalize this result below.

Proposition 2: Identification of Difference in Average $\beta\mu$-racial Bias. *If parallel disparities (Assumption 1) and parallel accurate statistical discrimination (Assumption 2) holds, the homophily estimand identifies the average difference in $\beta\mu$-racial bias between Black and white DMs. That is,*

$$\tau = \delta_W^{\beta\mu} - \delta_B^{\beta\mu}.$$

Proof. First, using Assumption 1 and Proposition 1, rewrite the homophily estimand as

$$\tau = \{E_{bw}[D(b) - D(w)]\} - \{E_{bb}[D(b) - D(w)]\}.$$

The above corresponds to the average difference in total racial bias across Black and white DMs. Next, substituting in the decomposition of the decision D, the two terms become:

$$E_{bw}[D(b) - D(w)] = E_{bw}[\widetilde{D}(b) + \widetilde{\beta\mu}(b) - \widetilde{D}(w) - \widetilde{\beta\mu}(w)]$$

$$E_{bb}[D(b) - D(w)] = E_{bb}[\widetilde{D}(b) + \widetilde{\beta\mu}(b) - \widetilde{D}(w) - \widetilde{\beta\mu}(w)].$$

Under PASD, the $\widetilde{D}$ terms cancel, leaving

$$\tau = E_{bw}[\widetilde{\beta\mu}(b) - \widetilde{\beta\mu}(w)] - E_{bb}[\widetilde{\beta\mu}(b) - \widetilde{\beta\mu}(w)] = \delta_W^{\beta\mu} - \delta_B^{\beta\mu}. \quad \square$$

Proposition 2 illustrates that the additional assumption of PASD result in the homophily estimand isolating differences in $\beta\mu$-racial discrimination rather than total racial discrimination.

Testing for the Presence of Bias. Under Assumptions 1 and 2, the homophily estimand can be used to test for the presence of $\beta\mu$-racial bias (or total racial bias under only assumption 1). The remark below summarizes several implications that affect the properties of such a test.

Remark 1 *Under Assumption 1 (and Assumption 2), the following are true.*

1. *Non-zero homophily ($\tau \neq 0$) implies that at least one case was affected by total ($\beta\mu$-)racial bias.*

2. *Positive homophily ($\tau > 0$), does not imply that there is only total ($\beta\mu$-)racial bias against Black filers, nor does $\tau < 0$ rule out some DMs exhibiting total ($\beta\mu$-)racial bias against Black filers.*

3. *Zero homophily ($\tau = 0$) does not imply that no cases are affected by total ($\beta\mu$-)racial bias, as this scenario could arise if there are DMs with opposing biases that cancel out on average.*

The above has direct implications for the ability to test for the presence of racial bias ($\delta^{ATT} \neq 0$ or $\delta^{\beta\mu} \neq 0$ for at least one case) by testing the null hypothesis of zero total bias on average ($H_0 : \tau = 0$). Such a test may fail to reject a null of zero racial bias even when some cases are affected by racial bias. However, a nonzero τ can only arise if racial bias affects at least one case. In this sense, the test is conservative. Formally, testing $\tau = 0$ is an underpowered test for the presence of bias but it still has exact size.

Quantifying the Impact of Bias. Our last result highlights what can be inferred about the scale of racial bias from an estimate of homophily. Without further assumptions, the homophily estimate can be used to partially identify average racial bias. We then show that the identified set for average racial bias can be further narrowed under additional assumptions, potentially yielding tight and informative lower bounds on the impact of racial bias on decisions. Our exposition here assumes both Assumption 1 (parallel disparities) and Assumption 2 (parallel accurate statistical discrimination) and hence focuses on quantifying $\beta\mu$-racial bias. Analogous results without Assumption 2 can be derived for total racial bias. The additional assumptions we consider are stated below.

Assumption 3 *On average, White DMs weakly exhibit bias against Black filers/in favor of non-Black filers:* $\delta_W^{\beta\mu} \geq 0$.

Assumption 4 *On average, Black DMs weakly exhibit bias against Black filers/in favor of non-Black filers:* $\delta_B^{\beta\mu} \geq 0$.

Note that for $\tau > 0$, Assumption 4 implies Assumption 3 (while the reverse is not true). Let $1 - p = \Pr(r_j = w)$, which corresponds to the proportion of white DMs. With this notation we can write $\delta^{\beta\mu} = p\delta_B^{\beta\mu} + (1-p)\delta_W^{\beta\mu}$. Note that neither assumption requires that all DMs of a given race exhibit the same direction of bias, rather, the assumptions relate to average bias among the groups of DMs.

The following proposition summarizes our partial identification results under Assumptions 1-4.

Proposition 3: Partial Identification of Average Bias *Suppose that homophily is positive ($\tau > 0$) and that parallel disparities (Assumption 1) and parallel accurate statistical discrimination (Assumption 2) both hold, then the statements below follow:*

1. *With no further assumptions, τ partially identifies $\delta^{\beta\mu}$ as follows: $\delta^{\beta\mu} \in [(1-p)\tau - 1, 1 - p\tau]$.*
2. *Under Assumption 3 ($\delta_W^{\beta\mu} \geq 0$), τ implies a higher lower bound, partially identifying $\delta^{\beta\mu}$ as follows: $\delta^{\beta\mu} \in [-p\tau, 1 - p\tau]$.*
3. *Under Assumption 4 ($\delta_B^{\beta\mu} \geq 0$), τ implies a higher lower bound, partially identifying $\delta^{\beta\mu}$ as follows: $\delta^{\beta\mu} \in [(1-p)\tau, 1 - p\tau]$.*

Proposition 3 shows how increasingly stronger assumptions allow researchers to obtain stricter lower bounds on the role of $\beta\mu$-racial bias in influencing dismissals ($\delta^{\beta\mu}$). Assumptions 1 and 2 imply an upper bound of $1 - p\tau$. With Assumption 4, the lower bound on the impact of $\beta\mu$-racial bias is the proportion of White DMs multiplied by the homophily estimand: $(1-p)\tau$. Another way to characterize the relative importance of $\beta\mu$-racial bias is to divide the identified set for $\delta^{\beta\mu}$ by the observed disparity $E[D|r_i = b] - E[D|r_i = w]$,which characterizes the share of the observed disparity due to $\beta\mu$-racial bias. Proposition 3 may aid future research on racial bias by helping researchers determine the appropriate identified set based on which assumptions they believe are appropriate for their setting. It also provides a "menu" of implications for the various sets of assumptions.

Our empirical analyses turns to estimating racial disparities in bankruptcy and homophily. Guided by the framework above, we test for violations of Assumptions 1 and 2, test for the presence of $\beta\mu$-racial bias, and report identified sets (using Proposition 3) to quantify racial bias.

Comparison with Related Models and Econometric Frameworks. Our bias detection and identification results build on a growing literature on the detection of bias. A central strand of this literature focuses on outcome tests (Becker, 1957, 1993; Canay et al., 2020; Hull, 2021). The key idea behind the outcome test is that bias can result in the marginal cost to the DM (e.g., default rates in lending) being lower among groups facing discrimination. Arnold, Dobbie and Yang (2018) propose an IV strategy, exploiting random assignment to DMs differing in their leniency, to overcome the challenge of estimating differences in outcomes for marginal agents from different groups. Their key insight is that the LATE identified by the IV can identify group-specific marginal costs.

The primary advantage of our homophily approach, compared to outcome tests, is that homophily does not require the researcher to observe (or know) the DM's objective function. Many economic and legal decisions have complex or abstract objectives, and sometimes additional outcomes such as DM compensation are also affected by their decision (as is the case with bankruptcy trustees). Homophily can therefore be used either as a complement to an outcome test analysis or in place of an outcome test when it is not feasible due to complexity or data limitations. A second appealing feature is that our framework suggests tests that can be used to falsify assumptions that impact whether the homophily estimand indicates $\beta\mu$-racial bias versus total racial bias (i.e., including accurate statistical discrimination).

Our framework most closely relates to that of Anwar and Fang (2006), which develops a decision model of police searches of motorists and identification results related to homophily. We generalize the decision model in several important dimensions. First, we allow DMs to value *multiple* outcomes and for DMs to differ in their preferences over outcomes. Second, our model allows for prediction errors, which may vary across DMs, and result in *inaccurate* statistical discrimination. Third, our model allows all outcomes (including decision costs) to depend flexibly on filer non-race characteristics that may be correlated with race.[12]

Our framework also has several econometric innovations relative to Anwar and Fang (2006). First, our test for the presence of bias will generally have more asymptotic power. Both tests may

[12]Anwar and Fang (2006) allows the benefit of the decision to flexibly depend on filer race and non-race characteristics, but the cost of the decision is allowed only to depend on the filer and DM's race. This would rule out, for example, police officers experiencing a higher cost of searching motorists with expensive vehicles (e.g., they may fear confronting wealthy/influential motorists).

fail to reject when bias is present. However, while our test will correctly reject the null of no bias in all cases when the Anwar and Fang (2006) test rejects, our test will also correctly reject in cases where the Anwar and Fang (2006) test fails to reject (asymptotically).[13] Our framework also differs in that we show how homophily can *quantify* bias by partially identifying the net share of dismissals due to bias. Lastly, our framework also emphasizes threats to identification (and proposes falsification tests) such as systematic differences along DM race in terms of how non-race characteristics influence their decisions.

3.3 Summary of Model Extensions

Our decision model can be extended in a variety of ways that make it more realistic without fundamentally changing our identification results. We discuss these extensions in detail in Appendix B and summarize them briefly here.

Noisy observation of x. We can generalize our setting to instead allow the DM to only observe a noisy signal related to the non-race characteristic x (rather than observing x directly). If DMs of different races tend to receive noisier signals or differ in their beliefs about how the signal predicts x, these can be sources of inaccurate statistical discrimination and would be reflected in the homophily estimate.

Dynamics and DM Learning. We consider a dynamic extension where the DM faces a sequence of cases and there may be spillovers from the outcomes in past cases to futures ones (which the DM may internalize). A natural spillover of interest is learning. In a dynamic setting, inexperienced DMs could have bigger prediction errors about how race and non-race characteristics predict outcomes Y (fraud, payment completion, etc.). Over time, DMs could learn about the relation between these variables and reduce prediction errors. If prediction errors tend to be largest when facing different-race filers, resulting in inaccurate statistical bias, learning could potentially reduce this form of bias over time. This extension does not significantly change the interpretation

[13] In our notation, Anwar and Fang (2006) rejects when $E_{bw}[D] - E_{bb}[D]$ and $E_{ww}[D] - E_{wb}[D]$ have opposite signs. For example, their test would reject if Black filers are more likely to be dismissed when facing white DMs and white filers are less likely to be dismissed when facing white DMs. However, suppose all filers are more likely to be dismissed when facing white DMs (making both of the differences positive). Anwar and Fang (2006) would fail to reject bias. However, if Black filers experience a *larger* increase in dismissal rates when moving from Black to white DMs (compared to white filers), our homophily estimand would still be positive and hence reject the null of no bias.

of homophily estimand. It highlights, however, that our homophily estimates average over many cases in a DM's career and will reflect the average influence of bias. Biases may change over time within DMs.

Allowing for Additional DM Actions. We augment the model to allow the DM to take multiple other actions. Specifically, the DM can take a number of actions during a "first stage" before choosing whether to dismiss in the "second stage." Notably, a stage 1 action could include choosing how much effort to exert to gather and process information. Allowing for other actions in a prior stage does not fundamentally change the identification results, but it helps to clarify the interpretation of the bias that homophily can detect and measure.

An interesting implication of this extension is that biases influencing stage 1 actions can also alter the impact of other biases on stage 2 (e.g., dismissal). For example, suppose an outcome that the DM cares about is preventing bankruptcy fraud, but she has taste-based bias that makes her dislike fraud more when it is committed by Black filers. Because of this bias, when she faces Blacker filers, she exerts more effort to gather and scrutinize information that could signal fraud (e.g., searching the filer's social media pages for evidence of extravagant spending). If the DM also overestimates the likelihood that Black filers commit fraud, confirmation bias could lead her beliefs to overreact to the additional information she collects. As a result, her decision to gather more information could further *reduce* the accuracy of her subjective expectation $E_j(\cdot)$, increasing the influence of inaccurate statistical bias on her subsequent dismissal decision. This extension highlights how biases, even different types, can compound in multi-stage decisions. Additionally, it illustrates how the bias that the homophily estimand detects with respect to the decision of interest (e.g., dismissal) may reflect DM biases from an earlier "stage" of the interaction.

Delegation. Lastly, we instead model the DM as making a "recommendation" to a third party who ultimately makes the decision. From the perspective of the DM, the actual decision becomes a function of her recommendation and "noise" introduced by the third party. Under this extension, the homophily estimand identifies the sum of "direct" bias present in the DM's recommendation and "indirect" bias from how the bias in their recommendation influences the "noise" (i.e., how bias in the trustee's recommendation alters the judge's decision). The indirect effect corresponds to the additional bias from the judge that is *induced* by the trustee's *biased* recommendation. For example, if a trustee discriminating makes a judge more willing to discriminate, the influence of the DM's bias on the judge's willingness to discriminate would also be reflected in our estimate. It also allows inaccurate statistical discrimination to "compound": when the trustee's recommendation reflects a biased forecast, a judge may also incorrectly update their beliefs, introducing more inaccurate statistical bias into the judge's ultimate decision. This highlights how a "chain" of biased DMs can amplify biases.

4 Data and Descriptive Facts

We now turn to the data used in our empirical analysis. The backbone of our dataset is court docket header information from the universe of personal bankruptcy filers in the Lexis Nexis Public Records database from 1990-2022. The filing header data includes the identity of the filer, trustee, and bankruptcy judge for a given bankruptcy proceeding. This data allows us to merge our bankruptcy cases with a dataset from the Federal Judicial Center (FJC) Integrated Data Base of all bankruptcy cases filed since Fiscal Year 2008. The FJC data has detailed information sourced from bankruptcy filings beyond the simple header information we observe for the universe of filings.

Panel A of Table 1 reports summary statistics for bankruptcy outcomes and characteristics. Dismissal and chapter status are observed for all 63 million cases in the Lexis Nexis data, while the other characteristics are observed only for the approximately 21 million cases that merge with the FJC data. The main reason for the drop off in the number of observations is the more limited time period covered by the FJC data.

Overall, 16% of bankruptcy cases are dismissed, meaning that the court terminated the case without allowing any debt relief. However, dismissal is virtually nonexistent for Chapter 7 (2% of cases) while it is the modal outcome for Chapter 13 cases (56% of cases). As discussed above, Chapter 13 cases involve payment plans; when debtors fail to adhere to the agreed-upon repayment plans, their cases are dismissed.

Turning to other characteristics of bankruptcy filers, 6% of petitioners file pro se, meaning they represent themselves instead of being represented by an attorney. About 14% of filers have filed before, especially among Chapter 13 filers (32%). Very few Chapter 7 filers report holding non-exempt assets, whereas almost all Chapter 13 filers report non-exempt assets. Roughly half of filers own a home, and roughly half are filing jointly with a co-petitioner (usually a spouse). The average petitioner has $3,750 in monthly income, $400,0000 in assets and 7 times as much debt as assets, with about half of their debt being secured. Chapter 7 petitioners anticipate still having $300 more in monthly expenses than income post-bankruptcy in contrast to Chapter 13 petitioners, who anticipate making $600 more in income each month than their expenses.

4.1 Race Imputation Algorithm

We impute the race of various parties (e.g., filers, trustees, judges, and attorneys) using a deep-learning model extending Kotova (2021). Our race imputation model predicts race using addresses (aggregated to the census block level) and full names. The algorithm employs both natural language processing (NLP) and a recurrent neural network (RNN) model.

We train our model using two datasets. The first is a large dataset containing names, addresses, and self-reported race for 36 million Americans: the universe of registered voters in seven states (Florida, North Carolina, South Carolina, Georgia, Tennessee, Louisiana, and Alabama). Florida and Georgia in particular both have relatively large Black and Hispanic populations, which should lead to more accurate and precise race imputation for minorities. The second dataset is US Census data on block, tract, zip code, and county racial composition for the 2010 and 2020 censuses.[14] We split full names into bigrams (all pairs of adjacent letters) and then apply RNN (using softmax activation) to the bigrams and the racial composition of the finest census geography available for the closest decennial census, training the model on the voter data to predict race using self-reported race.

We make several improvements over the original Kotova (2021) method. First, we allow the model to use the sequence of bigrams and not just the set.[15] Second, we use a more granular level of geography (census block) whenever possible; when data is unavailable for a census block, we use the racial composition of the smallest geography for which census data is available (either

[14] We use six variables to capture race composition: the fraction of residents identifying as Asian, Black, Hispanic, Other, or White and the fraction identifying as multiracial. In unreported testing results, we found model performance improvements when including the multiracial variable.

[15] We do so by using a one-hot encoding that preserves sequence order, which is not the case for the original Python function used in Kotova (2021).

census tract or ZIP code).[16] Third, instead of equally weighting the classification of each observation, which tends "prioritize" correctness for the majority class, we choose a set of weights that prioritizes average accuracy for *four* classes (Asian, Black, Hispanic, and White) and overall accuracy. By not also prioritizing "Other", our algorithm ends up classifying very few people as "Other". Effectively, the algorithm instead tries to best predict what race someone would self-report if they were asked to say something other than "Other". This feature has two advantages: (1) de-prioritizing "Other" allows us to improve the model's performance within the other four classes and (2) "encouraging" the algorithm to pick one of these four races may yield a measure of race that better corresponds to how other people would perceive a person's race.

To impute filer race, we input the filer's full name and local racial composition into the race imputation model. The model returns a predicted probability for the likelihood that the filer identifies as Asian, Black, Hispanic, Other, or White. Note that we therefore have a continuous measure for race. We generally use this continuous measure in our analysis (unless otherwise specified). On the one hand, our measure of race will generally be subject to measurement error, which would at worst attenuate our estimates if it is independent. However, a continuous measure may do a better job "measuring" race for bi-racial individuals.

For non-filers, such as judges, trustees, or attorneys, we do not observe their residential address in the bankruptcy court records. For now, we assume that their home residence is within the same MSA as their office address (for trustees and attorneys) or the same MSA as one of the division office (for judges). Judges often serve in more than one court division (e.g., the Alaska bankruptcy district has divisions in Anchorage, Fairbanks, Juneau, and Ketchikan). We treat all of the MSAs of the district's divisions as a geographical block within which we assume the judge resides somewhere. For example, the Honorable Arthur B. Briskman served for a time in the Florida Middle district. Within the Florida Middle District, there are four division offices (Ft. Myers, Jacksonville, Orlando, and Tampa). We assume that this judge resides somewhere in the union of the MSAs spanned by these four division offices. With this information, we can apply a less granular version of the algorithm described above. We are currently working to supplement addresses data

[16] We have six racial composition measures. Five are the shares of the population that identify as Asian, Black, Hispanic, Other, and White (either exclusively or in addition to other races). The sixth is the share of the population that identifies as more than one race.

for non-filers using data from Whitepages.com.[17]

4.2 Imputation Model Performance

We evaluate the performance of the model out-of-sample on a subset of 1 million voters excluded from the model training and hyper-parameter tuning. Table 2 reports several performance statistics by race. Overall, our model performs well; our imputation aligns with self-reported race 88% of the time. For Black individuals, our model performs similarly well, correctly identifying Black people 84% of the time. The algorithm performs better for larger race categories, but still obtains high levels of precision and recall for non-Other races (as discussed, our algorithm is optimized to perform best on the four major race categories at the expense of the Other category.) Figure 2 plots the Receiver Operator Characteristic (ROC) curves by race, showing that—again with the exception of the Other category—even at low false positive rates, the imputation algorithm achieves high true positive rates for each race. The figure also reports the area under the curve (AUC) for each ROC, which summarizes overall model performance in terms of how well predictions are ranked. The area under the curve (AUC) is 97% for Black people, which means that 97% of the time the model correctly ranks a random Black person as being more likely to be Black than a random non-Black person. The AUC is 97% for Asians, 98% for Hispanics, 95% for Whites, and 72% for "Other."

Our model outperforms other popular alternative imputation methods, including Bayesian Improved Surname Geocoding (BISG), which is used by regulators such as the CFPB, to test for fair lending compliance. Comparing the performance of BISG on the same subset of 1 million voters, we find an overall accuracy of 72%, but this method correctly identifies Black individuals only 51% of the time.[18] The performance of our model is closer the performance achieved by methods that use *images* of people to impute their race than to BISG. Greenwald et al. (2023) finds a correlation of 87% between their image-based and self-reported indicators for identifying as Black. Under this metric, BISG achieves a 52% correlation while our model achieves 80%. Figure 3 reports ROCs and AUCs for BISG. Our algorithm is performs better within every category under

[17] For judges, given their relatively low number, we also double-check these ethnicities by hand using internet searches.

[18] We use Surgeo's ZIP-level BISG procedure in this comparison. Using their "BIFSG," which also uses first names, leads to overall accuracy of 79% but the percent of Black people correctly identified remains low (49%), although the false positive rate for classifying people as Black somewhat reduces from 45% to 41% (it is 16% for our model).

these metrics. The performance differentials are largest for the Black, White, and Other classes. For example, BISG's AUC is 84–87% for Black people (versus our 97%) and 83–84% for White people (versus our 95%). For Other, BISG is approximately as accurate as random guessing.

4.3 Summary Statistics

Filer, Trustee, and Judge Race Composition. Panel B of Table 1 reports imputed race shares for our sample. We estimate that 74% of US bankruptcy filers are White, 14% are Black, and roughly equal shares of the remaining 12% are Asian, Hispanic, or Other. Comparing across chapters, Chapter 13 filers are twice as likely to be Black (23%) than Chapter 7 filers (11%).

Panel C of Table 1 reports the results of imputing trustee race. We estimate that 84% of US bankruptcy trustees are White, 8% are Black, and the remainder are roughly evenly split between Asians, Hispanics, and the other category. Although Chapter 13 trustees are slightly more likely to be Black than Chapter 7 trustees (10.4% vs. 7.6%), the distribution of trustee race is fairly similar across chapters.

Dismissal and Chapter Choice Disparities. Armed with imputed racial identity for each petitioner in the data, we calculate dismissal rates by racial group in Figure 1. Almost 35% of Black filers have their bankruptcy petitions dismissed compared with 23% for those identifying as "Other" , 19% for Hispanic filers, and 17% and 15% for Asian and White filers, respectively. One driver of racial disparities in dismissal rates is chapter choice. Black filers are the most likely to file under Chapter 13, at a rate of 44%. People identifying as "Other", White, or Hispanic file Chapter 13 in 25% of cases. Asian filers are the least likely to file under Chapter 13, doing so in 21% of cases.

If all filers had the same dismissal rates as White filers within that same chapter, overall Black dismissal rates would fall by over 10 percentage points. This indicates that within-Chapter racial disparities in dismissal rates account over half of the total Black-White disparity. Within-chapter disparities also contribute to the overall disparities for other non-White filers. Within-Chapter racial disparities in dismissals also account for a majority of the overall disparity for other non-White filers: 8pp for "Other", 4pp for Hispanic, and 4pp for Asian filers.

Chapter 13 Dismissal Disparities over Time. Figure 4 plots cumulative dismissal hazards by for Black and White Chapter 13 filers separately. For both races, the high dismissal rate for Chapter 13 cases represents a gradual increase in dismissals over time as debtors' repayment plans fail. Pairing petitioner and trustee imputed race measures, Figure 5 demonstrates that racial distribution of the assignment of trustees to filers in the data is nearly indistinguishable from the expected distribution under random assignment.

5 Results

5.1 Disparities in Personal Bankruptcy

Before testing for racial bias, we first document disparities across the race of the bankruptcy filer. In Tables 5 and 6, we examine what explains whether bankruptcy cases are more likely to be dismissed for Black filers. As discussed above, dismissal is equivalent to denying bankruptcy protection for these individuals and, as shown in Dobbie and Song (2015), case dismissal has severe negative consequences for the consumer including reducing earnings and increasing the likelihood of foreclosure and mortality. Table 5 focuses on Chapter 13 filers. Unconditionally, Black filers are 17 percentage points more likely to be dismissed relative to other filers. This estimate falls to 12 percentage points when including filer ZIP code effects in Column (3), suggesting that some of the disparity between races is related to factors associated with filers they live. Additional fixed effects continue to attenuate the coefficient estimates only slightly. Even in Column (5) with the full set of fixed effects, we still estimate that Black filers are 11 percentage points more likely to be dismissed from Chapter 13 after including all fixed effects. This is a 20% increase from the mean dismissal rate of 56% for these cases. Using the FJC controls in Column (6) attenuates the conditional Black-White dismissal gap, but this is mostly driven by the change in time period rather than an uneven distribution of the observable control variables across races. However, the coefficients on the controls in Column (6) help benchmark the economic magnitude of the racial disparity coefficient. The 9 percentage point effect is almost half of the large pro se effect of filing without professional legal counsel and two thirds of the magnitude of being a repeat filer.

Table 6 provides a useful contrast by examining predictors of dismissal for Chapter 7 filers. Depending on the controls, we estimate that Chapter 7 Black filers are between 0.4 and 3 percentage points more likely to be dismissed than other races. Relative to the average Chapter 7 dismissal probability of 2.4%, Black filers are about twice as likely to be dismissed, even when controlling for year, zip code, judge, and trustee fixed effects. Roughly half of the attenuation of the Black-White dismissal gap in Column (6) using controls in the FJC data is mostly driven by the necessary restriction of the Column (6) sample to 2009-2022, with the remainder being explained by the controls.

Taken together, Tables 6 and 5 show that Black filers are significantly more likely to be dismissed from both Chapter 7 and Chapter 13 bankruptcy, but the absolute size of the effect is an order of magnitude larger in Chapter 13. Contrasting the degree to which observable characteristics can explain racial disparities in dismissal rates across chapters, we note that the unexplainable portion of the racial gap is much larger in Chapter 13, where overall dismissal rates are higher and trustees have significantly more discretion in case outcomes. In the sections that follow, we more finely test how racial biases may be driving these outcomes.

5.2 How Much Does Bias Contribute to Disparities?

Table 7 tests whether filer-trustee homophily affects bankruptcy dismissal using specifications of the form

$$Dismissed_{igjkt} = \beta_0 BlackFiler_i + \tau BlackFiler_i \times WhiteTrustee_k + \alpha_t + \gamma_g + \delta_j + \mu_k + \varepsilon_{igjkt} \qquad (1)$$

where α_t, γ_z, δ_j, and μ_k are fixed effects for year, geography (county or zip code), judge, and trustee, respectively. Similar to Tables 6 and 5, the dependent variable is a dummy indicating the bankruptcy case was dismissed. $BlackFiler_i$ is the imputed probability that a filer's race is Black. Similarly, $WhiteTrustee_i$ is the imputed probability that a trustee's race is White. Judge and trustee fixed effects control for any fixed biases towards dismissal of the judge or the trustee.[19] These fixed effects also absorb the race of the judge and trustee such that we do not control separately for their race. As outlined in Section 3 above, we are interested in τ, which corresponds to the homophily estimand τ introduced in Section 3.2.1. Recall that that homophily estimand captures

[19]See Change and Schoar (2013), Dobbie and Song (2015), Bernstein, Colonnelli, and Iverson (2019) for evidence of fixed judge leniency tendencies. While we presume that trustees also exhibit biases, we are unaware of systematic evidence on this front.

how the difference in Black and non-Black filers changes when cases are assigned to a White trustee. Meanwhile, β_0 tests for differences in outcomes between Black and non-Black filers who are assigned to non-White trustees.

Table 7 reports the results of estimating (1) separately for Chapter 13 and Chapter 7 filings in Columns (1)-(2) and (3)-(4), respectively. For the full Lexis Nexis Chapter 13 sample in Column (1), the coefficient β_0 on the probability the petitioner is Black is 6.8 pp, and the interaction term Black Filer $\times$ White Trustee is 4.7 pp and significant at the 90% confidence level. Conditional on the controls available in the Lexis Nexis-FJC merged sample in Column (2), the Black filer coefficient is small and statistically insignificant, and the filer-trustee race interaction term is positive and statistically significant. Thus, Black filers have similar dismissal rates to non-Black filers when assigned to non-White trustees but are significantly less likely to receive full bankruptcy protection when assigned to a White trustee. As laid out in Section 3, since $\tau \neq 0$, we conclude that there is bias present in how trustees treat bankrupt consumers. Further, as long as we assume that trustees are as lenient towards filers of their own race as they are towards those of a different race, Table 7 is strong evidence of homophily effects among bankruptcy trustee, and this homophily effect is quite large.

Columns (3)-(4) repeat this exercise for Chapter 7 filers, where trustees play a more procedural and less discretionary role. The results show that there is essentially no homophily in Chapter 7 bankruptcy filings. Unconditionally, Black filers are 17 percentage points more likely to be dismissed from Chapter 13 than non-Black filers (Column (1) of Table 5). Thus, we estimate that about 44% of the overall disparity in Chapter 13 outcomes is due to racial bias among bankruptcy trustees. Importantly, Black filers are 77% more likely to use Chapter 13 bankruptcy—precisely the chapter of bankruptcy where bias is likely to work against them.

Returning to the framework of Section 3.2.1, we can bound racial bias using these homophily estimates. By Proposition 3, under Assumptions 1, 2, and 4, the total amount of racial bias $\delta^{\beta\mu}$ due to taste-based discrimination or inaccurate statistical discrimination is bounded by $(1-p)\tau$ and $1-p\tau$, where p is the White share of trustees and τ is the average difference in dismissal rates for Black filers facing White and non-White trustees. Using $p = 0.17$ and $\hat{\tau} = 0.074$, we estimate that the racial bias portion of the homophily coefficient τ is between 0.06 and 0.98. Comparing the lower bound with the total unconditional Black-non-Black disparity in Chapter 13 dismissal rates (17% in Table 5), we conclude that at least 35% of the racial disparity in Chapter 13 filing rates is due to trustee racial bias from either taste-based discrimination or inaccurate statistical discrimination. By contrast, the estimate of $\hat{\tau} = 0$ for Chapter 7 filers in Table 7 means that we have no evidence of racial bias among Chapter 7 trustees.

Association with Implicit Bias. Consistent racial bias contributing to the Black-White dismissal disparity, we further show that these disparities are correlated with a measure of implicit bias. We measure implicit bias using data from Project Implicit's Implicit Association Test (IAT). IAT scores correspond to a measure of subconscious bias. Experimental evaluations of the test have found that its measures of implicit bias are difficult for takers to manipulate (Banse et al., 2001; Egloff and Schmukle, 2002). Following Bursztyn et al. (2021), we focus on "forced" respondents who were required to take the test as either an assignment for work or school. We further restrict attention to White respondents. Our IAT sample spans 2010-2020 and we aggregate IAT scores to the county-year level. We similarly aggregate our case data to the county-year-chapter level and measure the average difference in dismissal rates between Black and White filers.

Figure 8 displays binscatters relating the Black-White dismissal disparity to IAT scores. Overall, we find a strong, positive association between implicit bias among White residents and the dismissal disparity for Chapter 13. A one standard deviation higher IAT score (indicating more anti-Black implicit bias) predicts approximately a 2 percentage point larger racial disparity in Chapter 13 dismissal rates. In contrast, we find nearly a perfectly flat relationship between implicit bias and Chapter 7 dismissal disparities, mirroring the disappearance of homophily in our Chapter 7 estimates. We stress that our IAT results reflect correlations and not causal estimates. The mirroring of our homophily estimates provides further suggestive evidence that bias (possibly implicit) contributes to the Black-White dismissal disparity.

5.3 Falsification Tests

As discussed in Section 3.2.1, the identification of the portion of the homophily coefficient due to racial bias relies on several assumptions. The parallel disparities and parallel accurate statistical discrimination (PASD) assumptions can be violated if filers can choose their DMs or vice versa, which could lead to unequal distributions of non-race characteristics X_i across DM race. In our setting, trustees are quasi-randomly assigned to filers, which should result in balanced case characteristics across trustees. Figure 5 shows that the actual pairing of filers to trustees by race is essentially identical to what would be expected under random assignment. Further, Figure 6 shows that filer characteristics do not meaningfully predict trustee race. This figure displays regression coefficients from a balance test where $WhiteTrustee_i$ is regressed on all filer characteristics, as well as bankruptcy chapter, year, zip code, and judge fixed effects. All coefficients are statistically indistinguishable from zero at the 95% confidence level. Most importantly, filer race is uncorrelated with trustee race. This alleviates concerns that the parallel disparities and PASD assumptions might be violated due to assortative matching between filers and trustees.

Another way that parallel disparities and PASD could be violated in our setting would be if White trustees respond to non-race characteristics differently than non-White trustees do. We test this formally by estimating our main homophily regression with augmented controls that interact $WhiteTrustee_i$ with case-level controls X_i. Because non-race characteristics systematically vary by race, it's possible that our homophily findings actually represent White trustees factoring in non-race characteristics into their dismissal decisions differently than non-White trustees. While such a parallel disparities violation would not necessarily mean that White trustees do not have racial bias, such a finding would cloud the interpretation of homophily as racial bias as it would complicate assigning differences in dismissal rates by trustee-filer race pairings to race per se.

We plot the coefficients on these interaction terms separately by chapter in Figure 7. Focusing on the Chapter 13 results first, the top row plots the interaction term of the probability a trustee is White and the probability a filer is Black, replicating our main homophily coefficient in the presence of controls interacted with trustee race.[20] Even controlling for eight interaction terms with trustee race, the coefficient on filer race is essentially unchanged from our main specification and is statistically significant at the 90% confidence level. With limited exceptions, the remaining coefficients are statistically insignificant. Although the confidence intervals are often wide—power is an issue when estimating all of these interaction terms simultaneously—the coefficients estimates

[20]The differential response of White trustees to case characteristics for Chapter 7 filings are very small and, for almost all characteristics, statistically insignificant.

themselves are generally small. The interaction terms are statistically significant at the 95% confidence level for only one characteristic: White trustees are less likely to lower-income filers than non-White trustees are. If Black filers have higher income on average, then this lower sensitivity to income by White trustees could explain some of the homophily effect. However, if the main homophily coefficient were actually picking up differential response to income, we would expect this coefficient to shrink once we account for these interactions. The inability of conditioning on these interaction terms to explain away the main homophily coefficient suggests that these characteristics and White trustee sensitivities do not vary significantly enough with race to violate parallel disparities or PASD. Overall, the results in Figure 7 are consistent with the homophily estimand identifying racial bias.

5.4 Mechanisms

The contrast between the share of Chapter 13 and 7 dismissal rate disparities attributable to trustee racial bias is consistent with an important role played by trustee discretion, which is substantial for Chapter 13 and nearly non-existent for Chapter 7. One of the possible places where any explicit or implicit bias by trustees may affect outcomes is in the determination of Chapter 13 plan payments. Following a Chapter 13 filing, the trustee makes a determination about what level of debt payments to creditors a Chapter 13 petitioner can afford. Planned payments are heavily influenced by an estimate of the filer's disposable income, defined as the gap between forecasted income and allowable expenses. If trustees forecast higher income or allow fewer expenses, trustees could reach a determination of higher disposable income and therefore more onerous creditor payment plans. Using data on income-expense gaps and the identification strategy leveraging quasi-random assignment of trustees to filers, we can test whether there is a homophily effect in income-expense gaps paralleling the racial homophily in dismissals documented above.

Unconditionally, we find that Black filers have a lower income-expense gap, roughly \$163 lower among Chapter 13 filers. However, for Black filers quasi-randomly assigned to White trustees, the income-expense gap is \$120 higher, suggesting higher planned payments that filers would have a harder time completing. Moreover, the homophily result for the income-expense gap holds even conditional on income, suggesting that White trustees are especially strict with

Black filers in their allowable expenses. Combined with the frequency of case dismissals for failure to pay planned payments, the disproportionate share of Black petitioner filings that are Chapter 13, and the growth in the racial dismissal gap over time, this provides evidence that racial bias in discretionary judgments by a key gatekeeper in the personal bankruptcy system play an important role in explaining why Black personal bankruptcy filers are far less successful in obtaining debt relief.

6 Conclusion

An extensive literature documents racial disparities in household finance. In this paper, we provide direct evidence of racial disparities in bankruptcy outcomes, a system that provides more debt relief each year than the resources allocated by all state unemployment insurance programs combined. We leverage a new dataset built on the near universe of US personal bankruptcy filings from 2010-2022, augmented with deep-learning imputed measures of race using names and addresses. Comparing the dismissal rates for filers imputed to be Black against the dismissal rates of filers from all other races, our most conservative estimate controlling for a rich set of non-race case characteristics is a 33% (10 percentage point) higher dismissal rate for Black filers than non-Black Chapter 13 filers. Looking across chapters, chapter choice plays a large role in explaining higher dismissal rates by Black filers, with non-race observable characteristics explaining most of Chapter 7 dismissal disparities by race but on half of the unconditional Chapter 13 dismissal rate gap by race.

After developing new results formalizing how homophily can detect and quantify bias—even in settings where outcome tests are infeasible—we provide novel evidence on the importance of bankruptcy trustees as important and discretionary intermediaries in the bankruptcy process. Black filers filing for Chapter 13 are more likely to have their cases dismissed when randomly assigned to a White trustee, which can explain 70% of the residual disparity in overall dismissal rates between Black and White filers. Additional evidence suggests that White trustees exercise their discretion in part by allowing lower expenses by Black filers, leading to requiring higher required payments to creditors. Interpreting these estimates using our decision-model framework, our estimates are consistent with trustees' taste-based or inaccurate statistical discrimination. We conclude that bias among bankruptcy decision-makers may significantly limit Black households' access to debt relief.

These results suggest several practical policy implications worthy of further investigation. Most directly, the homophily results suggest that increased diversity among legal decisionmakers could help close the racial gap in bankruptcy outcomes by improving the odds that minority filers are paired with minority trustees. Second, our results suggest that collecting data on filer races could be an important first step towards transparency and accountability, similar to the intent of the Home Mortgage Disclosure Act's collection and disclosure of data on protected class membership for mortgage applicants. Third, the sharp contrast between almost no racial gap in outcomes for Chapter 7 and substantial disparities for Chapter 13 filings—combined the high overall failure rate of Chapter 13 bankruptcy plans and other evidence about the costliness of filing for Chapter 13—could motivate efforts to reconsider the utility of Chapter 13 altogether or to reduce the barriers to filing for Chapter 7, such as reducing upfront legal costs and or increasing means tests thresholds for Chapter 7 to make them bind for fewer filers. Finally, the apparent role of trustee discretion suggests that effective trustee training or process reform could improve outcomes for Black bankruptcy filers. Implicit bias training could reduce unconscious bias. Standardizing or automating more of the filing processing could limit the scope for inaccurate statistical bias to affect outcomes.

References

Albanesi, Stefania and Jaromir Nosal, "Insolvency After the 2005 Bankruptcy Reform," 2018.

Anwar, Shamena and Hanming Fang, "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence," *American Economic Review*, 2006, *96* (1), 127–151.

_ , Patrick Bayer, and Randi Hjalmarsson, "The impact of jury race in criminal trials," *The Quarterly Journal of Economics*, 2012, *127* (2), 1017–1055.

_ , _ , and _ , "A jury of her peers: The impact of the first female jurors on criminal convictions," *The Economic Journal*, 2019, *129* (618), 603–650.

_ , _ , and _ , "Politics in the courtroom: Political ideology and jury decision making," *Journal of the European Economic Association*, 2019, *17* (3), 834–875.

Argyle, Bronson, Benjamin Iverson, Taylor Nadauld, and Christopher Palmer, "Personal Bankruptcy, Moral Hazard, and Shadow Debt," 2020. Working Paper.

Arnold, David, Will Dobbie, and Crystal S Yang, "Racial bias in bail decisions," *The Quarterly Journal of Economics*, 2018, *133* (4), 1885–1932.

_ , _ , and Peter Hull, "Measuring Racial Discrimination in Bail Decisions," April 2020. NBER Working Paper No. 26999.

Auclert, Adrien, Will S. Dobbie, and Paul Goldsmith-Pinkham, "Macroeconomic Effects of Debt Relief: Consumer Bankruptcy Protections in the Great Recession," 2019.

Banse, Rainer, Jan Seise, and Nikola Zerbes, "Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT," *Zeitschrift für experimentelle Psychologie*, 2001, *48* (2), 145–160.

Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace, "Consumer-Lending Discrimination in the FinTech Era," June 2019. NBER Working Paper No. 25943.

Bayer, Patrick, Fernando Ferreira, and Stephen L Ross, "What drives racial and ethnic differences in high-cost mortgages? The role of high-risk lenders," *The Review of Financial Studies*, 2018, *31* (1), 175–205.

Becker, Gary S., *The Economics of Discrimination*, University of Chicago Press, 1957.

_ , "Nobel Lecture: The Economic Way of Looking at Behavior," *Journal of Political Economy*, 1993, *101* (3), 385–409.

Begley, Taylor and Amiyatosh Purnanandam, "Color and Credit: Race, Regulation, and the Quality of Financial Services," *Journal of Financial Economics*, Forthcoming.

Blattner, Laura and Scott Nelson, "How Costly Is Noise? Data and Disparities in the US Mortgage Market," 2020.

Bohren, Aislinn, Kareem Haggag, Alex Imas, and Devin Pope, "Inaccurate Statistical Discrimination: An Identification Problem," *NBER Working paper No. 25935*, 2020.

Bohren, J Aislinn, Peter Hull, and Alex Imas, "Systemic discrimination: Theory and measurement," 2022.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer, "Stereotypes," *Quarterly Journal of Economics*, 2016, *131*, 1753–1794.

Braucher, Jean, Dov Cohen, and Robert M Lawless, "Race, attorney influence, and bankruptcy chapter choice," *Journal of Empirical Legal Studies*, 2012, *9* (3), 393–429.

Bursztyn, Leonardo, Thomas Chaney, Tarek A Hassan, and Aakaash Rao, "The immigrant next door: Exposure, prejudice, and altruism," *NBER Working paper*, 2021, (w28448).

Butler, Alexander W, Erik J Mayer, and James P Weston, "Racial disparities in the auto loan market," *The Review of Financial Studies*, 2023, *36* (1), 1–41.

Canay, Ivan A., Magne Mogstad, and Jack Mountjoy, "On the Use of Outcome Tests for Detecting Bias in Decision Making," *NBER working paper*, 2020, (w27802).

Chatterjee, Satyajit, Dean Corbae, Makoto Nakajima, and José-Víctor Ríos-Rull, "A quantitative theory of unsecured consumer credit with risk of default," *Econometrica*, 2007, *75* (6), 1525–1589.

Dávila, Eduardo, "Using elasticities to derive optimal bankruptcy exemptions," *The Review of Economic Studies*, 2020, *87* (2), 870–913.

Dobbie, Will and Jae Song, "Debt relief and debtor outcomes: Measuring the effects of consumer bankruptcy protection," *American Economic Review*, 2015, *105* (3), 1272–1311.

_, Paul Goldsmith-Pinkham, and Crystal S. Yang, "Consumer Bankruptcy and Financial Health," *Review of Economics and Statistics*, 2017, *99* (5), 853–869.

_, _, Neale Mahoney, and Jae Song, "Bad Credit, No Problem? Credit and Labor Market Consequences of Bad Credit Reports," *Journal of Finance*, 2020.

Egloff, Boris and Stefan C Schmukle, "Predictive validity of an Implicit Association Test for assessing anxiety.," *Journal of personality and social psychology*, 2002, *83* (6), 1441.

Elias, Stephen and Leon Bayer, *The new bankruptcy: will it work for you?*, Nolo, 2017.

Frame, W Scott, Ruidi Huang, Erik J Mayer, and Adi Sunderam, "The impact of minority representation at mortgage lenders," Technical Report, National Bureau of Economic Research 2022.

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther, "Predictably unequal? the effects of machine learning on credit markets," *Journal of Finance*, Forthcoming.

Ganong, Peter, Damon Jones, Pascal Noel, Fiona Greig, Diana Farrell, and Chris Wheat, "Wealth, race, and consumption smoothing of typical income shocks," 2020. NBER Working Paper No. 27552.

Ghent, Andra C, Ruben Hernandez-Murillo, and Michael T Owyang, "Differences in subprime loan pricing across races and neighborhoods," *Regional Science and Urban Economics*, 2014, *48*, 199–215.

Greenwald, Daniel, Sabrina T Howell, Cangyuan Li, and Emmanuel Yimfor, "Regulatory Arbitrage or Random Errors? Implications of Race Prediction Algorithms in Fair Lending Analysis," *Working Paper*, 2023.

Gross, Tal, Matthew J. Notowidigdo, and Jialan Wang, "The Marginal Propensity to Consume over the Business Cycle," *American Economic Journal: Macroeconomics*, 2020, *12* (2), 351–84.

Herkenhoff, Kyle, Gordon Phillips, and Ethan Cohen-Cole, "How Credit Constraints Impact Job Finding Rates, Sorting & Aggregate Output," 2019.

Hull, Peter, "What Marginal Outcome Tests Can Tell Us About Racially Biased Decision-Making," 2021.

Imai, Kosuke and Kabir Khanna, "Improving ecological inference by predicting individual ethnicity from voter registration records," *Political Analysis*, 2016, pp. 263–272.

Indarte, Sasha, "Moral Hazard versus Liquidity in Household Bankruptcy," *Journal of Finance (forthcoming)*, 2023.

Iverson, Benjamin, "Get in line: Chapter 11 restructuring in crowded bankruptcy courts," *Management Science*, 2018, *64* (11), 5370–5394.

_ , Joshua Madsen, Wei Wang, and Qiping Xu, "Financial Costs of Judicial Inexperience," 2020. Working Paper.

Jiang, Erica Xuewei, Yeonjoon Lee, and Will Shuo Liu, "Disparities in consumer credit: The role of loan officers in the FinTech era," 2021.

Keys, Benjamin J., "The Credit Market Consequences of Job Displacement," *The Review of Economics and Statistics*, 2018, *100* (3), 405–415.

Kiel, Paul and Hannah Fresques, "Data Analysis: Bankruptcy and Race in America," Technical Report, Pro Publica 2017.

Knowles, John, Nicola Persico, and Petra Todd, "Racial Bias in Motor Vehicle Searches: Theory and Evidence," *Journal of Political Economy*, 2001, *109* (1), 203–229.

Kotova, Nadia, "A Deep Learning Approach to Predicting Race Using Personal Name and Location (Natural Language Processing)," 2021.

Lefgren, Lars and Frank McIntyre, "Explaining the puzzle of cross-state differences in bankruptcy rates," *The Journal of Law and Economics*, 2009, *52* (2), 367–393.

_ , Frank L McIntyre, and Michelle Miller, "Chapter 7 or 13: Are client or lawyer interests paramount?," *The BE Journal of Economic Analysis & Policy*, 2010, *10* (1).

Livshits, Igor, James MacGee, and Michele Tertilt, "Consumer bankruptcy: A fresh start," *American Economic Review*, 2007, *97* (1), 402–418.

Morrison, Edward R, Belisa Pang, and Jonathon Zytnick, "Manipulating Random Assignment: Evidence From Consumer Bankruptcies in the Nation's Largest Cities," *Columbia Law and Economics Working Paper*, 2019, (614).

Morse, Adair and Karen Pence, "Technological Innovation and Discrimination in Household Finance," February 2020. NBER Working Paper No. 26739.

Munnell, Alicia H, Geoffrey MB Tootell, Lynn E Browne, and James McEneaney, "Mortgage lending in Boston: Interpreting HMDA data," *The American Economic Review*, 1996, pp. 25–53.

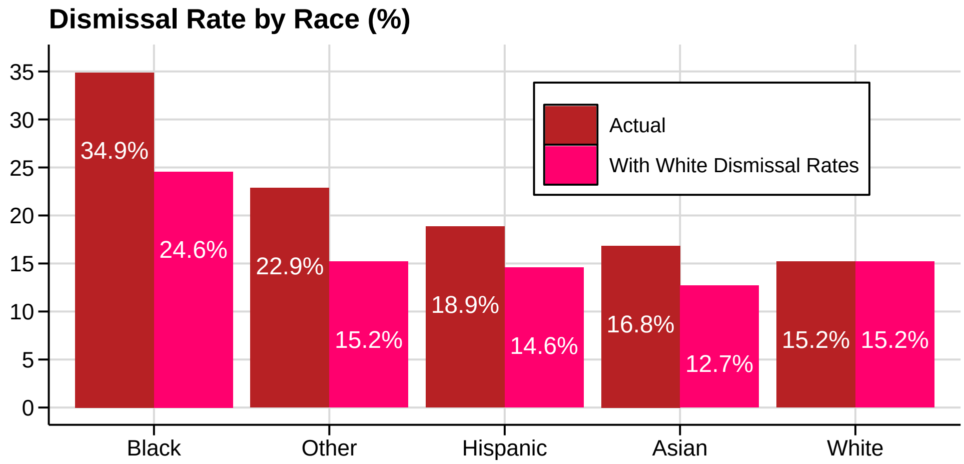
Musto, David K., "What Happens When Information Leaves a Market? Evidence from Post-bankruptcy Consumers," *The Journal of Business*, 2004, *77* (4), 725–748.

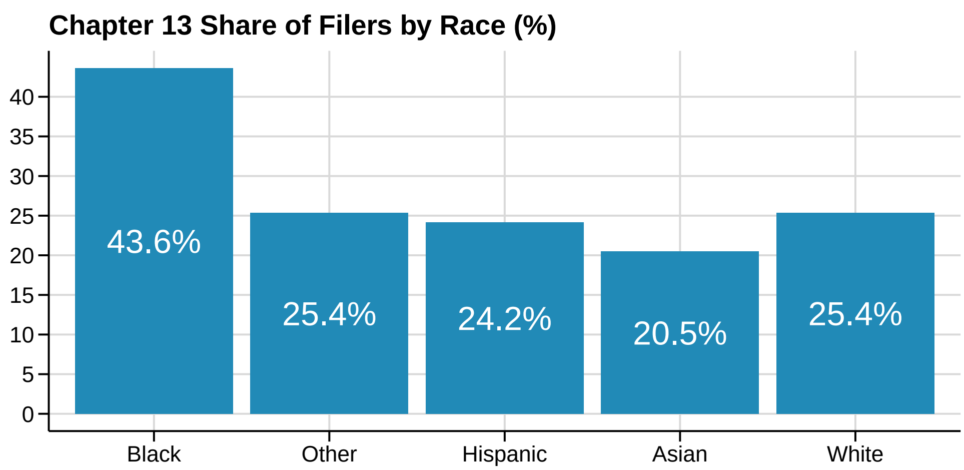
Reid, Carolina K, Debbie Bocian, Wei Li, and Roberto G Quercia, "Revisiting the subprime crisis: The dual mortgage market and mortgage defaults by race and ethnicity," *Journal of Urban Affairs*, 2017, *39* (4), 469–487.

Spriggs, William, "Is now a teachable moment for economists?," 2020. Federal Reserve Bank of Minneapolis.

Stavins, Joanna, "Credit card borrowing, delinquency, and personal bankruptcy," *New England Economic Review*, 2000, (Jul), 15–30.

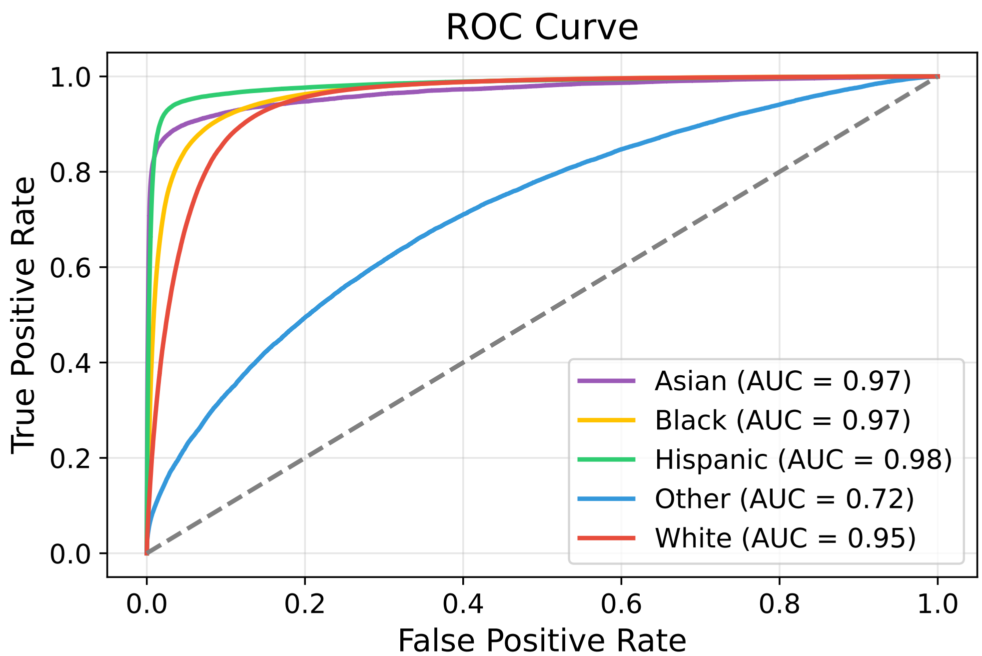
Figure 1: Bankruptcy Dismissal Rates and Chapter Choice by Petitioner Race





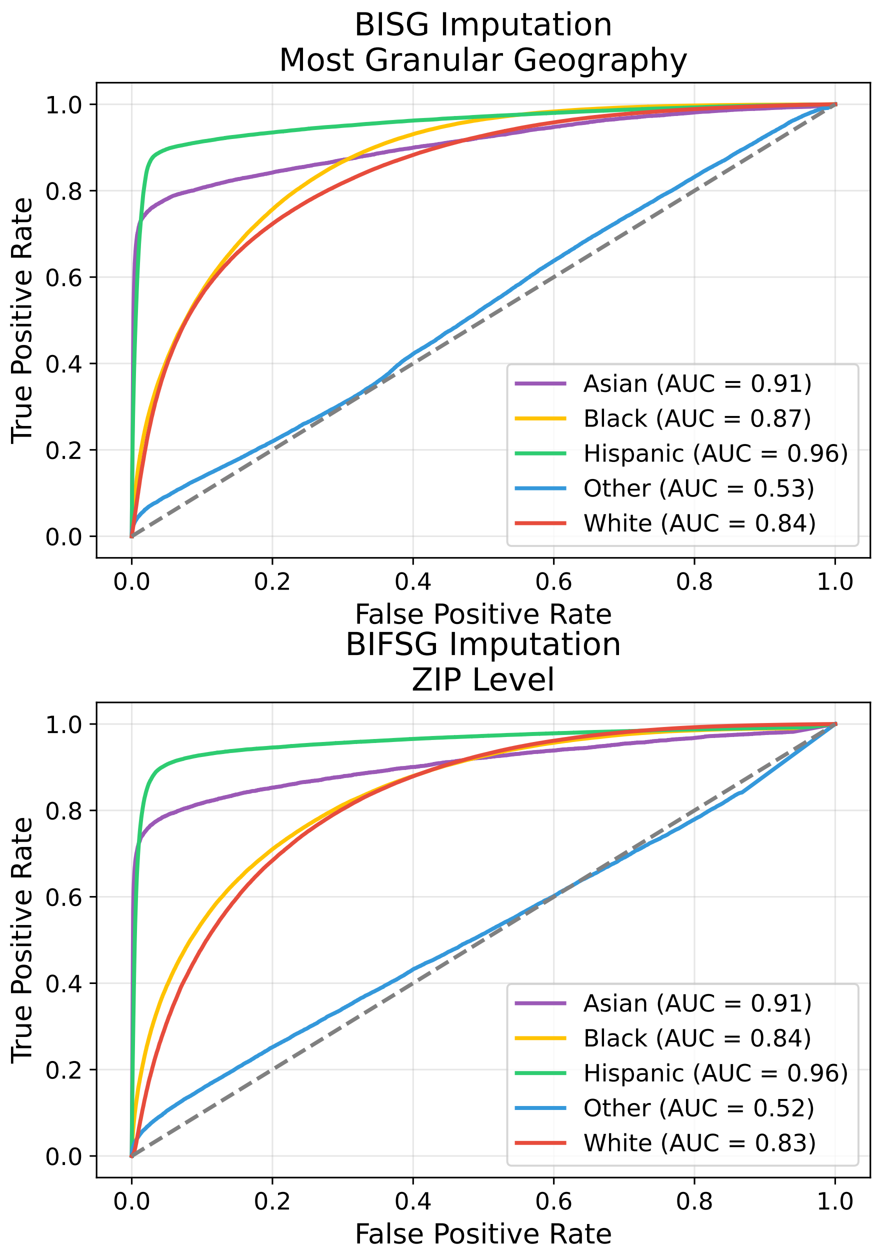
Notes: This figure plots bankruptcy dismissal rates and Chapter choice by filer race. The top graph reports dismissal rates for filers under either Chapter (7 or 13). The bottom graph reports the share of cases filed under Chapter 13. In the top graph, the dark red bars report the *actual* dismissal rate in our sample. The lighter pink bars next to them are *counterfactual* dismissal rates that help decompose the role of Chapter choice in accounting for overall disparities. The counterfactual holds constant each filer's Chapter choice but instead gives each filer the same *within-Chapter* dismissal rate as White filers. The counterfactual therefore reports dismissal disparities that would arise if they were only driven by Chapter choice. The difference between actual and counterfactual dismissal rates are attributable to *within-Chapter* racial disparities in dismissal rates. Filer race is imputed using the deep-learning algorithm described in Section 4.1.

Figure 2: Race Imputation Model Performance (ROC Curves)



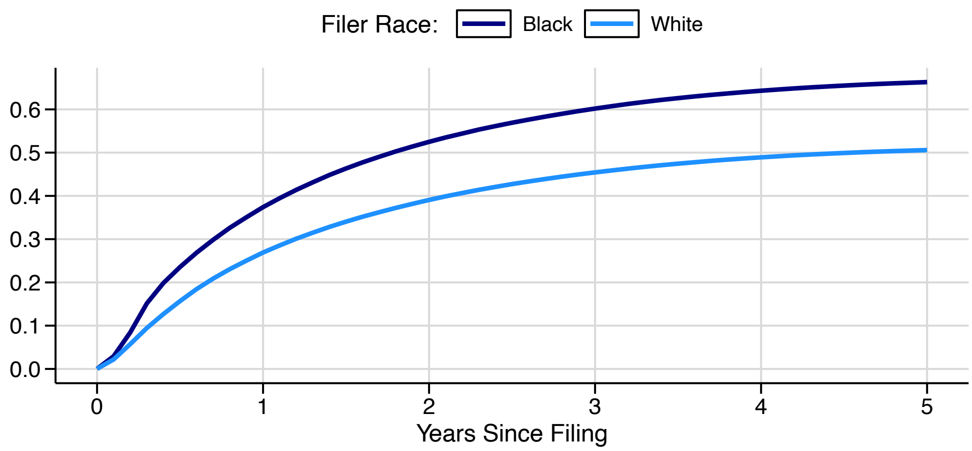
Notes: This figure plots a measure of model performance for our race imputation algorithm described in Section 4.1. The Receiver Operating Characteristic (ROC) curves report the true positive rate (recall) attained for a given tolerance of false positives. The true positive rate is the ratio of true positives to the sum of true positives and false negatives. The false positive rate is the ratio of false positives to false positives and true negatives. These rates are calculated using a one million observation, out-of-sample subset (out of sample in the sense that they are not used in the model training nor hyper-parameter tuning). The fast rise of the ROC curves indicates that our classification algorithm needs to tolerate very few false negatives in order to obtain true positive rates near 90% (except for "Other"). We also plot the 45-degree line as a performance benchmark; an algorithm that is equally likely to be correct or wrong would result in such a line. The AUC statistics for each race measure the area under the ROC curve, which equals 1 under perfect classification and 0.5 under random classification.

Figure 3: BISG and BIFSG Model Performance (ROC Curves)



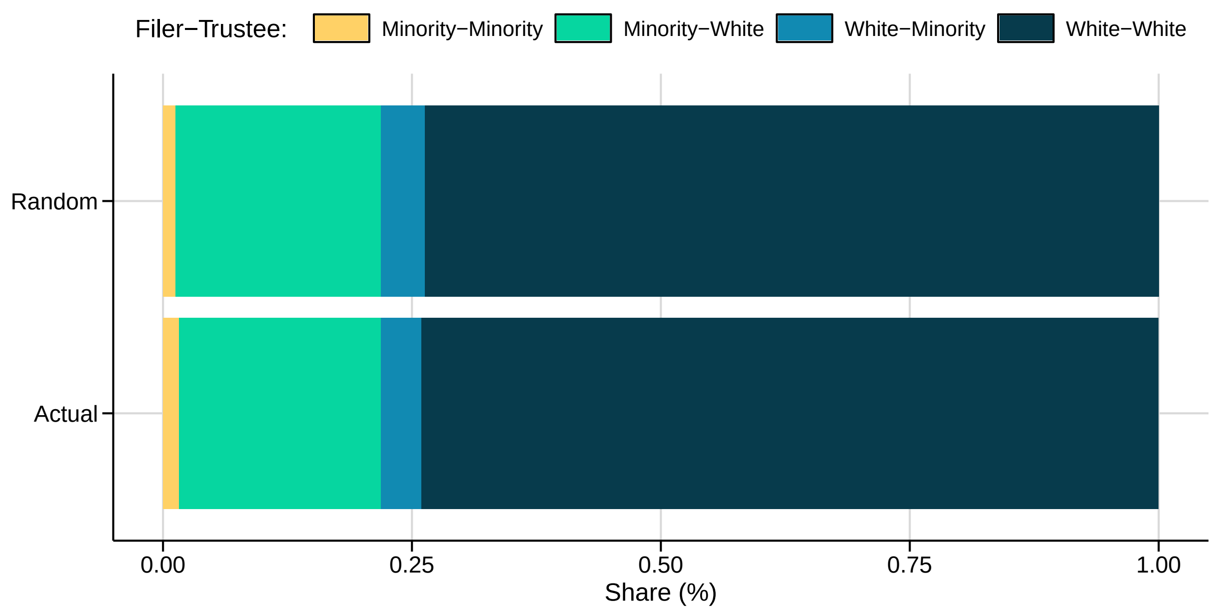
Notes: The figures plots ROC curves for alternative, popular race imputation algorithms. The top plot displays performance results for BISG (described in the main text) using Census tract whenever possible (576,472 observations) and ZIP code otherwise (357,350 observations). 66,178 observations are dropped because their surname does not appear in the Census database used to implement BISG. The bottom figure reports results from BIFSG (which uses first name in addition to surname) and ZIP code. We implement both algorithms using the Python package Surgeo. Overall, our algorithm outperforms BISG and BIFSG, especially for Black and White people.

Figure 4: Chapter 13 Bankruptcy Dismissal Survival Curves by Petitioner Race



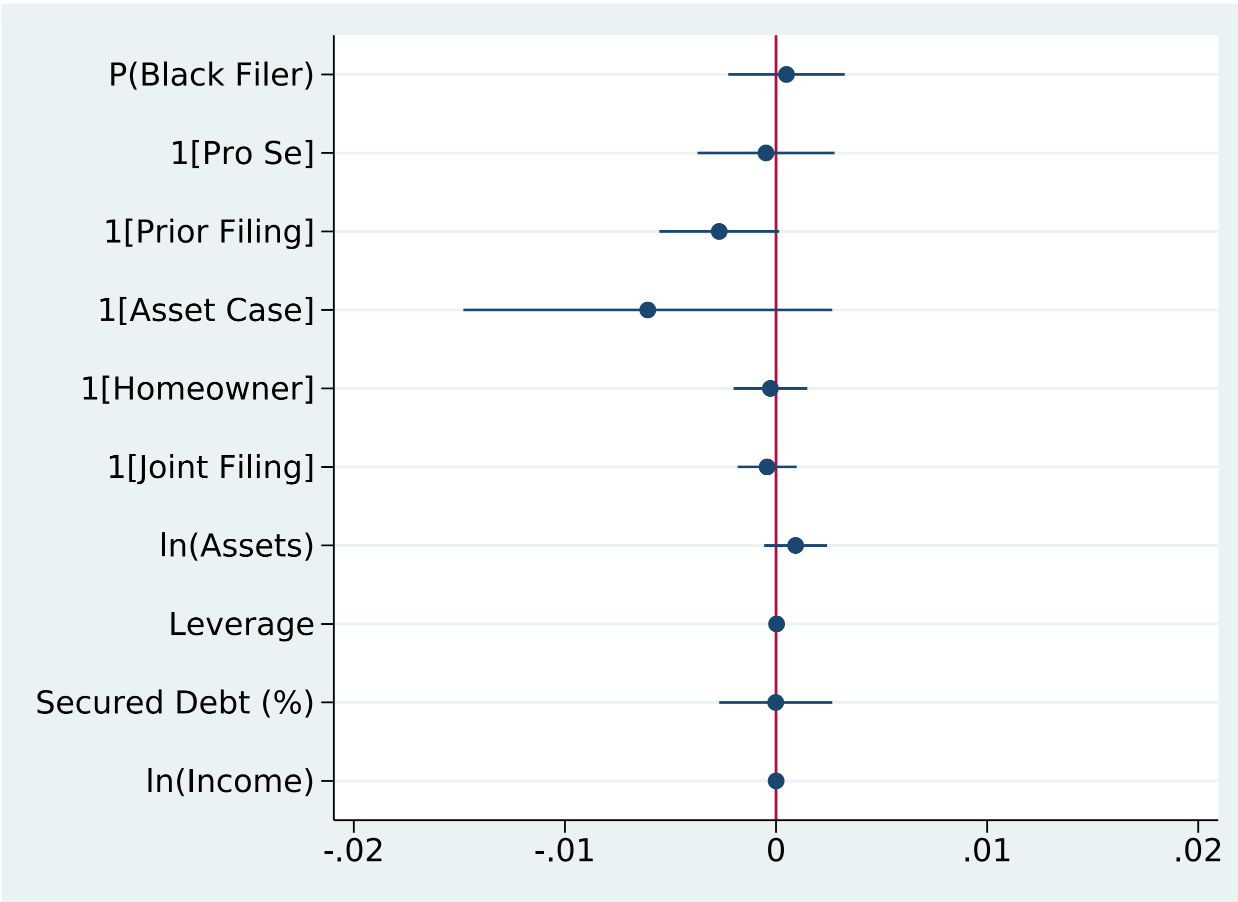
Notes: Figure plots cumulative dismissal hazard curves by race for Chapter 13 filers. The dark and light blue lines show the total fraction of Black and White Chapter 13 filers, respectively, that have had their cases dismissed within the indicated number of years since their initial filing.

Figure 5: Personal Bankruptcy Filings by Match with Petitioner Race



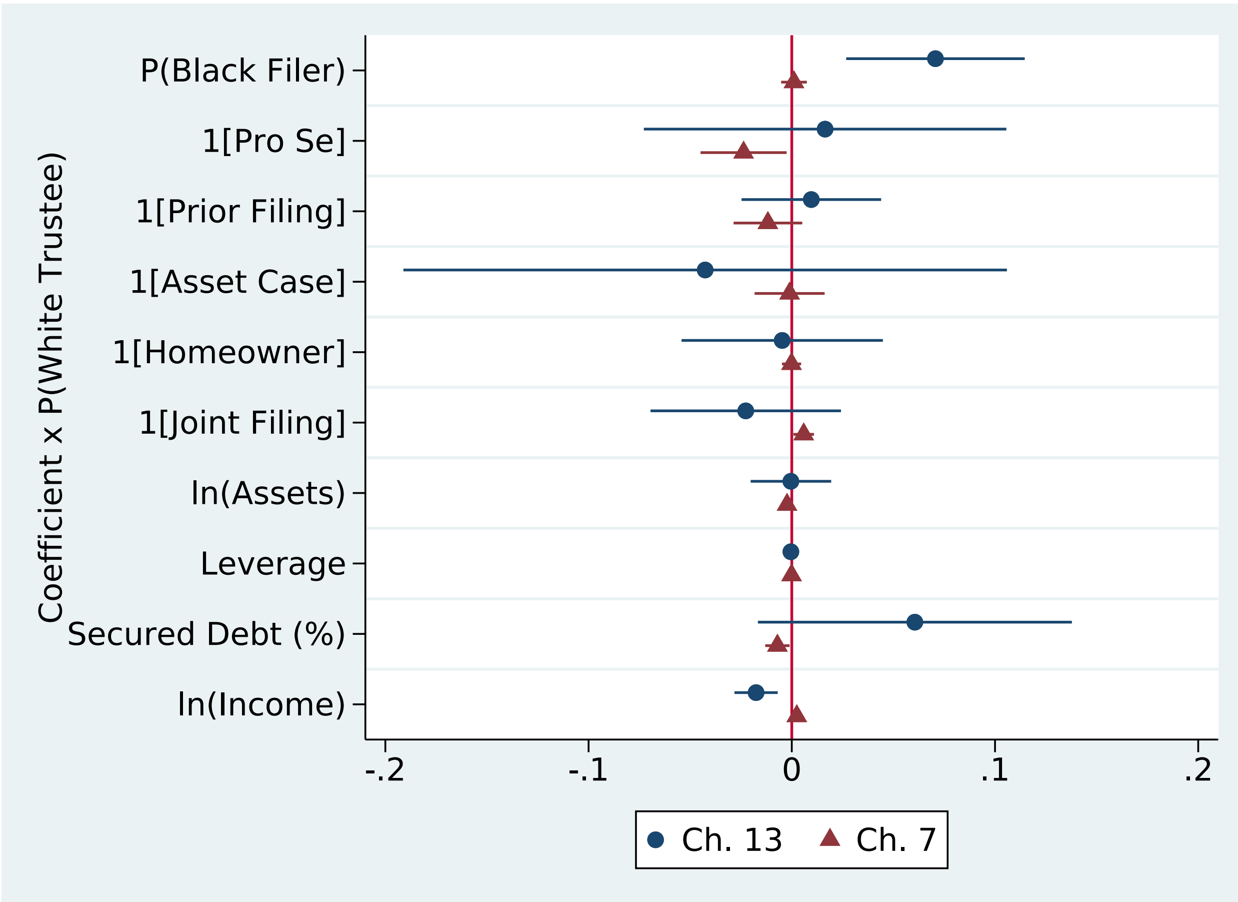
Notes: Figure plots share of cases where the bankruptcy petitioner's race matches with the race of each type of other participant in the bankruptcy proceeding. Bars labeled random report the share of matches that would belong to each race pair category if matching were random nationwide.

Figure 6: Testing for Correlation between Filer Characteristics and Trustee Race



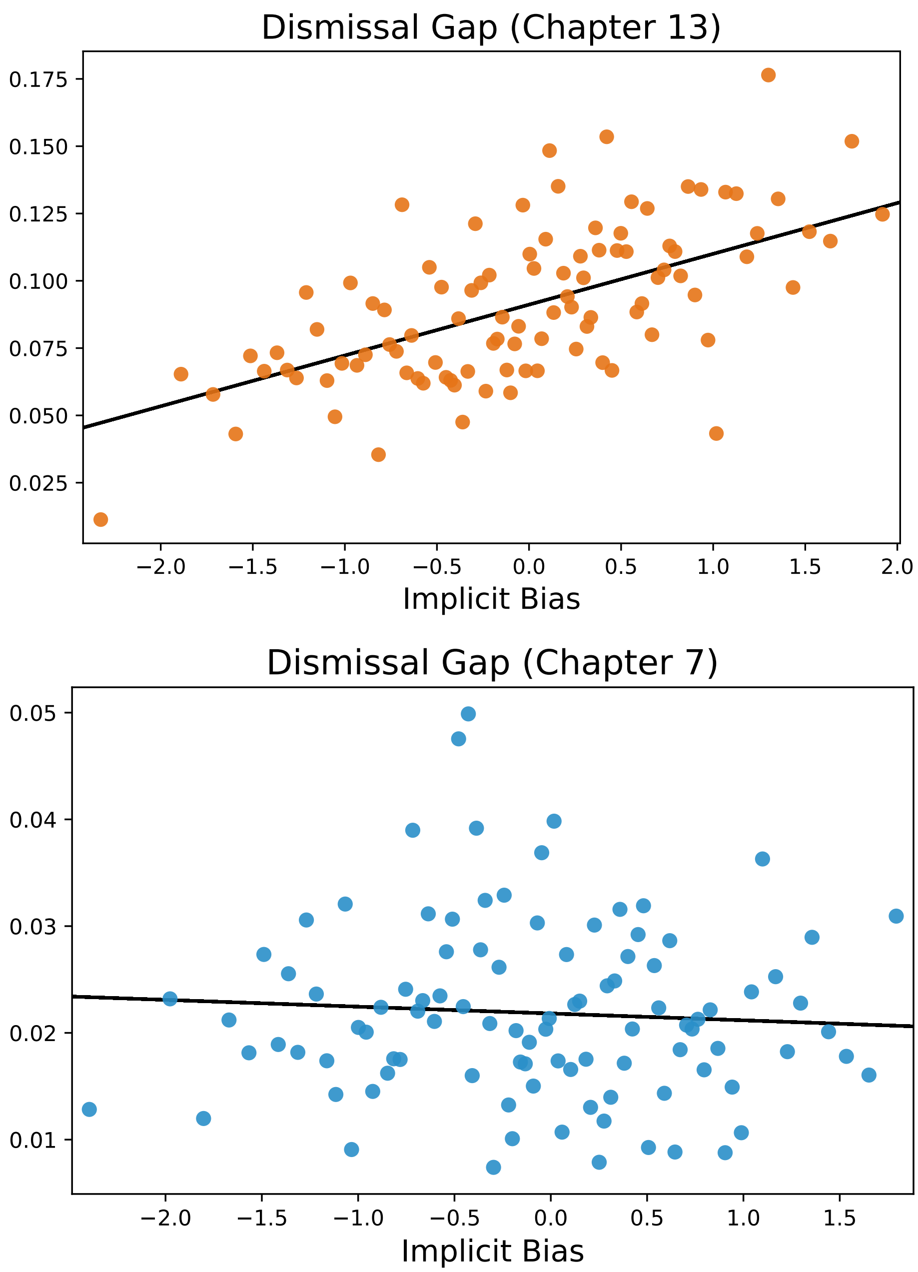
Notes: Figure plots coefficients from a regression testing for balance of filer characteristics by trustee race. The dependent variable is $WhiteTrustee_i$, and the coefficient for each independent variable is shown in the figure. The regression also includes fixed effects for filing year, zip code, and judge, and a Chapter 7 indicator. Standard errors are clustered by zip code and trustee.

Figure 7: Testing for Differential White Trustee Sensitivity to Non-Race Characteristics



Notes: Figure plots coefficients on interaction terms between the probability a trustee is white and several bankruptcy filing characteristics along with 95% confidence intervals clustered at the ZIP level. For each characteristic, blue dots report the interaction coefficient for Chapter 13, and the orange triangles report the Chapter 7 results.

Figure 8: Implicit Bias versus Black-White Dismissal Disparities



Notes: Figure depicts binscatter plots correlating the Black-White dismissal disparity against a measure of implicit bias. Our implicit bias measure comes from Project Implicit's Implicit Association Test (IAT). IAT scores for White respondents are averaged at the county-year level and then de-meaned and scaled by their standard deviation for ease of interpretation. A higher IAT scores corresponds to higher anti-Black implicit bias.

Table 1: Descriptive Statistics on Personal Bankruptcy Petitioners

	Chapters 7 and 13		Chapter 7		Chapter 13	
	Mean	N	Mean	N	Mean	N
Panel A: Bankruptcy Outcomes						
Dismissal	0.16	63,210,223	0.02	46,559,929	0.56	16,650,294
Filed Ch. 7	0.74	63,210,223	1	46,559,929	0	16,650,294
Pro Se	0.06	20,502,247	0.05	14,549,435	0.07	5,952,812
Prior Filer	0.14	20,248,920	0.07	14,491,525	0.32	5,757,395
Has Nonex. Assets	0.35	20,478,568	0.08	14,532,686	0.99	5,945,882
Owns Home	0.55	20,643,958	0.52	14,320,242	0.60	6,323,716
Joint Filing	0.45	21,554,305	0.44	14,702,321	0.48	6,851,984
Assets ($000s)	400.02	21,554,295	423.60	14,702,314	349.41	6,851,981
Debt/Assets	7.25	19,464,969	8.40	14,157,684	4.18	5,307,285
Secured Debt (%)	0.48	19,384,807	0.43	14,106,056	0.63	5,278,751
Monthly Inc. ($000s)	3.76	20,673,915	3.82	14,335,967	3.62	6,337,948
Monthly Inc. - Exp.	-1.78	20,661,108	-265.94	14,326,987	595.73	6,334,121
Panel B: Filer Race						
Asian	0.020	53,125,258	0.021	39,002,506	0.016	14,122,752
Black	0.142	53,125,258	0.112	39,002,506	0.227	14,122,752
Hispanic	0.056	53,125,258	0.058	39,002,506	0.052	14,122,752
White	0.742	53,125,258	0.769	39,002,506	0.665	14,122,752
Other	0.040	53,125,258	0.040	39,002,506	0.041	14,122,752
Panel C: Trustee Race						
Asian	0.010	58,566,649	0.011	43,058,405	0.005	15,508,244
Black	0.083	58,566,649	0.076	43,058,405	0.104	15,508,244
Hispanic	0.024	58,566,649	0.025	43,058,405	0.021	15,508,244
White	0.839	58,566,649	0.843	43,058,405	0.829	15,508,244
Other	0.044	58,566,649	0.045	43,058,405	0.042	15,508,244
Panel D: Judge Race						
White	0.81	1,247,291	0.82	922,948	0.79	324,343
Black	0.10	1,247,291	0.10	922,948	0.11	324,343
Hispanic	0.04	1,247,291	0.04	922,948	0.05	324,343
Asian	0.02	1,247,291	0.02	922,948	0.02	324,343
Other	0.02	1,247,291	0.02	922,948	0.02	324,343
Panel E: Attorney Race						
White	0.70	112,933	0.69	87,291	0.70	25,642
Black	0.09	112,933	0.09	87,291	0.08	25,642
Hispanic	0.17	112,933	0.17	87,291	0.17	25,642
Asian	0.02	112,933	0.02	87,291	0.03	25,642
Other	0.02	112,933	0.02	87,291	0.02	25,642

Notes: Table reports summary statistics for bankruptcy outcomes (panel A) and imputed race measures for filers, trustees, judges, and attorneys in panels B-E, respectively.

Table 2: Racial Imputation Performance Statistics

Race	Precision	Recall	F1-Score	Support
Asian	0.68	0.81	0.74	17,423
Black	0.84	0.84	0.84	237,780
Hispanic	0.83	0.89	0.86	74,649
Other	0.73	0.02	0.04	34,417
White	0.91	0.94	0.92	635,721
Macro Avg	0.80	.70	0.69	999,990

Notes: Table reports racial imputation performance statistics by race using the classification algorithm described in section 4.1. Precision is the share of true positives among all imputed positives. Recall is the share true positives among all actual positives. F1-score is the harmonic mean of precision and recall. The overall accuracy of the imputation (fraction correctly predicted) is 88%. We calculate our performance statistics on a one million observation, out-of-sample subset (i.e., this subset is set aside from the full L2 sample and not used to train the model nor tune its hyperparameters). 10 observations are dropped because neither their block, tract, nor ZIP can be linked to our Census race composition data.

Table 3: BISG Race Imputation Performance Statistics

Race	Precision	Recall	F1-Score	Support
Asian	0.74	0.57	0.64	14,836
Black	0.70	0.45	0.55	226,422
Hispanic	0.79	0.79	0.79	67,312
Other	0.53	0.01	0.03	31,175
White	0.77	0.92	0.84	594,077
Macro Avg	0.71	0.55	0.57	933,822

Notes: Table reports race imputation performance statistics for the Surgeo BISG algorithm using the most granular possible geography (576,472 at the tract level, and 357,350 at the ZIP level). Precision is the share of true positives among all imputed positives. Recall is the share true positives among all actual positives. F1-score is the harmonic mean of precision and recall. The overall accuracy of the imputation (fraction correctly predicted) is 76%. We use the sample subsample in the table evaluating our model's performance. More observations are dropped here because BISG requires that the surname be present in the database containing the racial composition associated with a given surname.

Table 4: BIFSG Race Imputation Performance Statistics

Race	Precision	Recall	F1-Score	Support
Asian	0.77	0.56	0.65	9,016
Black	0.59	0.49	0.54	150,949
Hispanic	0.80	0.78	0.79	51,928
Other	0.31	0.02	0.03	24,221
White	0.83	0.90	0.86	552,062
Macro Avg	0.66	0.55	0.57	788,176

Notes: Table reports race imputation performance statistics for the Surgeo BIFSG algorithm at the ZIP code level. Precision is the share of true positives among all imputed positives. Recall is the share true positives among all actual positives. F1-score is the harmonic mean of precision and recall. The overall accuracy of the imputation (fraction correctly predicted) is 79%. We use the sample subsample in the table evaluating our model's performance. Even more observations are dropped here because BIFSG requires that both the first name and surname be present in the database containing the racial composition associated with names.

Table 5: Dismissal Effects by Petitioner Race: Chapter 13

	(1)	(2)	(3)	(4)	(5)	(6)
Black Filer	0.171***	0.160***	0.118***	0.112***	0.110***	0.089***
	(0.015)	(0.014)	(0.005)	(0.005)	(0.004)	(0.004)
Pro Se						0.225***
						(0.011)
Prior Filer						0.136***
						(0.003)
Nonex. Assets						-0.052***
						(0.0123)
Owns Home						-0.049***
						(0.004)
Joint Filing						-0.078***
						(0.003)
ln(Assets)						-0.018***
						(0.002)
Debt/Assets						0.001***
						(0.0001)
Secured Debt (%)						0.209***
						(0.009)
ln(Monthly Inc.)						-0.020***
						(0.001)
Constant	0.559***					
	(0.010)					
N	6,667,799	6,667,798	5,517,052	5,371,214	5,370,748	2,593,043
R^2	0.016	0.042	0.227	0.258	0.277	0.257
Year FE		✓	✓	✓	✓	✓
Filer ZIP FE			✓	✓	✓	✓
Judge FE				✓	✓	✓
Trustee FE					✓	✓

Notes: Table reports regressions of an indicator for whether a Chapter 13 bankruptcy petition was dismissed in court onto the imputed probability the filer's race is Black. Control variables include indicator variables for whether filing was conducted without an attorney (*Pro Se)*, if the individual has filed a bankruptcy case in the previous 8 years (*Prior Filer*), if the filing has non-exempt assets that can be distributed to creditors (*Nonex. Assets*), if the individual is a homeowner (Owns Home), and if the filing was a joint filing with a spouse or domestic partner (*Joint Filing*). Continuous control variables are the log of total assets (*ln(Asset*s)), the total debt-to-asset ratio winsorized at the 1% level (*Debt/Assets*), the share of total debt that is secured (*Secured Debt (%)*), the log of monthly income (*ln(Monthly Income)*), and the difference between a filers monthly income and expenses winsorized at the 1% level (*Monthly Inc. - Exp.*). Robust standard errors are clustered at the ZIP code and trustee level and are displayed in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Dismissal Effects by Petitioner Race: Chapter 7

	(1)	(2)	(3)	(4)	(5)	(6)
Black Filer	0.029***	0.028***	0.025***	0.025***	0.025***	0.004***
	(0.002)	(0.002)	(0.0009)	(0.0009)	(0.0009)	(0.001)
Pro Se						0.098***
						(0.002)
Prior Filer						0.042***
						(0.001)
Nonex. Assets						-0.003***
						(0.001)
Owns Home						-0.001***
						(0.0003)
Joint Filing						-0.003***
						(0.0003)
ln(Assets)						0.001***
						(0.0002)
Debt/Assets						0.0001***
						(0.0000)
Secured Debt (%)						0.012***
						(0.001)
ln(Monthly Inc.)						-0.003***
						(0.0001)
Constant	0.024***					
	(0.0006)					
N	18,219,599	18,219,597	14,507,556	13,910,832	13,910,493	7,300,083
R^2	0.002	0.005	0.124	0.119	0.124	0.053
Year FE		✓	✓	✓	✓	✓
Filer ZIP FE			✓	✓	✓	✓
Judge FE				✓	✓	✓
Trustee FE					✓	✓

Notes: Table reports regressions of an indicator for whether a Chapter 7 bankruptcy petition was dismissed in court onto the imputed probability the filer's race is Black. Control variables include indicator variables for whether filing was conducted without an attorney (*Pro Se)*, if the individual has filed a bankruptcy case in the previous 8 years (*Prior Filer*), if the filing has non-exempt assets that can be distributed to creditors (*Nonex. Assets*), if the individual is a homeowner (Owns Home), and if the filing was a joint filing with a spouse or domestic partner (*Joint Filing*). Continuous control variables are the log of total assets (*ln(Asset*s)), the total debt-to-asset ratio winsorized at the 1% level (*Debt/Assets*), the share of total debt that is secured (*Secured Debt (%)*), the log of monthly income (*ln(Monthly Income)*), and the difference between a filers monthly income and expenses winsorized at the 1% level (*Monthly Inc. - Exp.*). Robust standard errors are clustered at the ZIP level and are displayed in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Dismissal Effects by Trustee and Petitioner Race

	(1)	(2)	(3)	(4)
Sample	Chapter 13		Chapter 7	
Black Filer	0.068***	0.025	0.029***	0.005*
	(0.024)	(0.020)	(0.005)	(0.003)
Black Filer × White Trustee	0.047*	0.074***	-0.004	-0.0007
	(0.028)	(0.024)	(0.006)	(0.003)
N	4,193,355	2,044,884	11,126,421	6,004,449
R^2	0.278	0.256	0.117	0.052
Controls		✓		✓
Disposition Year FE	✓	✓	✓	✓
Filer ZIP FE	✓	✓	✓	✓
Judge FE	✓	✓	✓	✓
Trustee FE	✓	✓	✓	✓

Notes: Table reports effects of filer and trustee race on an indicator for whether the bankruptcy petition was dismissed in court for Chapter 13 filings and Chapter 7 filings in columns (1)-(2) and (3)-(4), respectively. Black filer is the imputed probability the filer's race is Black. White trustee is an indictor for whether the court-appointed trustee's race is imputed to be White. Controls include all variables discussed in Table 6, and columns with controls use the FJC sample. Robust standard errors are two-way clustered at the ZIP code and trustee levels and are displayed in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Online Appendix

A Additional Figures and Tables

B Theory Appendix

B.1 Decision Model Extension: Noisy Observation of Non-Race Characteristic

One way to enrich the model is to suppose that, instead of observing the filer's non-race characteristic x, the DM observes a noisy signal s that may be correlated with x. In the context of bankruptcy, x could be the filer's ability to complete a Chapter 13 plan, which is ex ante uncertain to the DM. In this scenario, the filer's recent income history could be s.

For ease of exposition, suppose that the DM observes

$$s = x + \sigma.$$

Where σ is a random variable that may be correlated with x,r_i, or j. The DM may have inaccurate beliefs about the distribution of σ (including how it is correlated with other variables). These inaccurate beliefs, as well as the noise introduced by σ, can both be sources of prediction error.

The DM's payoff, $\Delta(j, r_i, x)$, still depends on x, but now x is ex ante uncertain. When making her decision, her subjective expectation now conditions on s instead of x:

$$D(j, r_i, x) = 1\left\{E_j\left[\Delta(j, r_i, x)|r_i, s\right]\right\}.$$

In presence of this model feature, we generalize the definition of prediction errors to reflect errors due to both inaccurate beliefs ($E(\cdot)$ vs $E_j(\cdot)$) and due to the imperfect observation of x (x vs s):

$$\mu(j, r_i, x, s) \equiv E\left[\Delta(j, r_i, x)|r_i, x\right] - E_j\left[\Delta(j, r_i, x)|r_i, s\right].$$

The prediction error now depends on both the signal s and the true x. As before, inaccurate statistical discrimination arises when changing only the filer's race changes the prediction error:

$$\mu(j, r_i = b, x, s) \neq \mu(j, r_i = w, x, s),$$

equivalently:

$$E\left[\Delta(j, r_i, x)|r_i = b, x\right] - E_j\left[\Delta(j, r_i, x)|r_i = b, s\right] \neq E\left[\Delta(j, r_i, x)|r_i = w, x\right] - E_j\left[\Delta(j, r_i, x)|r_i = w, s\right].$$

With a signal instead of a perfectly observed x, what changes relative to our baseline model is that there is now an additional channel through which inaccurate statistical discrimination may differ across DMs. Because the signal can flexibly depend on j, r_i, x, it is possible, for example, that White DMs tend to receive noisier signals about x when facing a Black filer. This extension underscores how imperfect information can contribute to or exacerbate inaccurate statistical discrimination.

For this extension, our definitions for taste-based and accurate statistical discrimination remain unchanged; we continue to define those objects with respect to accurate conditional expectations (i.e., conditioning on x).

B.2 Decision Model Extension: Dynamics and DM Learning

The model in the main text is static in the sense that there is no explicit time dimension. One feature of bankruptcy, and other potential settings of interest, is that DMs may learn information about a filer over time. For example, in Chapter 13 cases, trustees may observe how the filer's income evolves over the 5-year repayment plan. To allow for this, we can interpret x as the DM's information set at the time that they are making their decision, without explicitly modeling the evolution of their information set.

We can also add a further dynamic interpretation to our model by allowing the vector of outcomes Y to include outcomes related to future cases. For example, the DM could anticipate that her decision to dismiss i's case may affect her likelihood of dismissing future cases, detecting fraud in the future, or the predictability of how their future choices relate to distant outcomes. Regarding the latter example, it is in this sense that our model could feature DM learning over time, as well as sophistication in the form of predicting how their actions affect their future information sets.

B.3 Decision Model Extension: Additional DM Actions

In our model, the DM optimizes by choosing one variable: whether to dismiss. In reality, DMs may take a number of actions that may also influence their dismissal decision. For example, a trustee may choose how much time to invest in reviewing the filer's submitted paperwork.

A natural way to incorporate this feature into our setting is to break the DM's decision into two stages. The second stage consists of the dismissal decision (as modeled in the main text). The first stage entails the DM choosing a vector of actions $a \in A(j, i)$, where their choice set $A(j, i)$ may depend on the identity of the DM and the filer that they face (and, hence, the filer's race and non-race characteristics). The stage 1 actions could affect the vector of outcomes Y (e.g., exerting more effort could increase total non-leisure time spent on the case) and the DM's information set (e.g., the DM may be able to exert effort to improve precision in predicting the relationship between r_i, x, and Y). A sophisticated DM could choose their stage 1 actions while considering how these choices will affect their stage 2 dismissal decision.

The main impact of the multi-stage extension would be to allow for interesting interactions between statistical and taste-based discrimination. For example, suppose the DM's sense of "fairness" is one component of the outcome Y that they care about. Taste for discrimination could manifest as a DM only valuing fairness in cases with White filers. This, in turn, could lead the DM to invest less effort in stage 1 to improve the accuracy of their subjective expectation $E_j(\cdot)$ when facing a Black filer. In this sense, taste-based bias can result in inaccurate statistical discrimination.

B.4 Decision Model Extension: Delegation

Below we augment the model to instead feature a DM that makes a "recommendation." The fundamental change from our baseline model is that now the DM does not directly choose the case outcome but instead only influences it. We will further suppose that the econometrician only observes the case outcome D but not the DM's recommendation R. We write dismissal as a function of the DM's recommendation and a random variable e:

$$D(j, r_i, x, R, e) = R(j, r_i, x) + e(j, r_i, x, R).$$

From the perspective of the DM, it is as if their decision is subject to “noise” e. We allow the “noise” e to be correlated with j, r_i, x, R. Above we write e as a (stochastic) function of these objects. Moreover, we allow for the DM to have incorrect beliefs about the statistical relationship between noise e and j, r_i, x, R.

In the context of our bankruptcy application, we can interpret the DM as the trustee and the judge as a source of “noise”. We highlight the key implications of this extension below.