Journal of Law and Courts (2024), 1–22
doi:10.1017/jlc.2024.2

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Justice Isn't Blind: Attorney Attractiveness and Success in US Federal Court*

Nicholas W. Waterbury

University of Illinois at Springfield†
Email: nwate3@uis.edu

(Received 18 July 2023; Revised 30 November 2023; Accepted 07 February 2024)

Abstract

Many court observers have argued that judges are capable of avoiding the use of extraneous factors in decision-making. This study examines the influence of seemingly irrelevant heuristics on real-world courtroom proceedings. Drawing on theories from neighboring disciplines, I hypothesize that physically attractive attorneys have greater success in US federal court. Using a generalizable causal inference strategy and a dataset of over 1,000 cases and 3,000 votes, I find support for my expectations using multiple measures of attractiveness. These findings raise serious normative concerns about equality and underscore the need to adjust traditional models of judicial behavior to account for inherent biases.

Keywords: judicial behavior; political psychology; decision-making; attractiveness

As he sat on the bench listening to oral arguments, Supreme Court Justice Harry Blackmun often took notes on the day's proceedings. The notes contained legal analysis, observations on the behavior of his colleagues, and some surprising evaluations of the attorneys arguing before him. "Plump," "hairpiece with double-chin," and "young, sandy, nice looking," were among the descriptions Blackmun used to describe the physical appearance of lawyers.[1] Although Blackmun was perhaps the bluntest, he is hardly the

*The author thanks Joshua Boston, Lee Epstein, Pauline Kim, William Lowry, Andrew Martin, Jessica Schoenherr, Alan Simmons, James Spriggs, Amy Steigerwalt, Justin Wedeking, and three anonymous reviewers for their valuable suggestions and comments in the creation of this article. Previous versions of this study were presented at the 2021 annual meeting of the American Political Science Association, at the 2021 New Directions in Law and Society Workshop at the University of Massachusetts Amherst, and at a 2022 meeting of the Washington University in St. Louis American Politics Workshop. Replication materials are available at the Journal of Law and Courts Dataverse: https://doi.org/10.7910/DVN/OVMN95.

†Assistant Research Director, Center for State Policy and Leadership, PAC 464, One University Place, Springfield, IL 62703; nwate3@uis.edu.

[1]The notes from these cases, *Illinois* v. *Krull* 1986, *Board of School Commissioners* v. *Jacobs* 1974, and *Rogers* v. *Loether* 1973, respectively, can be viewed at the The Digital Archives of Justices Blackmun and

only jurist to have offered comments on attorneys' appearances. His colleague, Justice Antonin Scalia, advised male attorneys that they should never appear in court with "long hair in a ponytail" (Scalia and Garner 2008, 162), and US Court of Appeals Judge Ruggero Aldisert went a step further and wrote that an attorney's appearance was a critical part of their attempt to persuade judges (Aldisert 2010, 355). Based on these accounts, it seems clear that judges take note of the way attorneys look, but can positive (or negative) impressions actually affect the way judges decide cases? Put another way, do better-looking attorneys have an empirical advantage over their less attractive peers?

This paper examines one potential instance of heuristics influencing judicial decision-making by testing the effect of an attorney's attractiveness on their success in court. Dating from pioneering work (e.g., Dion, Berscheid and Walster 1972) to the present (e.g., Olivola, Tingley and Todorov 2018), a variety of researchers have found that attractive individuals enjoy benefits their colleagues do not. Humans are instinctively more likely to associate an attractive individual with other positive qualities, such as a high work ethic, friendly demeanor, or intelligence. As a result, attractive people enjoy many advantages in life, including employment decisions (Gilmore, Beehr and Love 1986), attention from medical doctors (Nordholm 1980), and grades in school (DeMeis and Turner 1978). Given that this effect influences decision-makers in all walks of life, judges are likely to similarly give advantages to attractive individuals in court. In the world of appellate judicial decision-making, I theorize that attractive attorneys are implicitly associated with positive qualities in the eyes of a judge, such as ideological agreement or sound legal reasoning. Attractive attorneys are therefore more likely to receive an individual judge's vote at the conclusion of a case. I also expect that the attractiveness advantage is universal and is given to attorneys by all of the judges on the panel tasked with deciding the case on the merits. Thus, attractive attorneys will also be more likely to win a case as a whole than less attractive attorneys.

To best identify the causal effect of attractiveness on judicial behavior, I draw comparisons between attorneys that are equal, save their attractiveness. I fielded a survey in which respondents rated the attractiveness of attorneys based on professional images obtained from the attorneys' websites. The sample of attorneys included all of those that orally argued cases at the US Courts of Appeals from 2017 to 2019. I then used contemporary matching algorithms to create a dataset of identical attorneys save the "treatment" variable (attractiveness). The matched attorneys were then compared based on their performance against a common opponent, the US government. With a final dataset containing more than 3,000 judicial votes and 1,000 cases, I find that more attractive attorneys are consistently and significantly more likely to win individual judges' votes and cases than their less attractive counterparts. This effect holds even when accounting for traditional predictors of success like ideology, experience, and financial resources. I replicate this analysis with alternative attractiveness ratings from a machine learning algorithm, a second survey based on the attorney's appearance in courtroom video recordings, and based on an attorney's rating relative to their opponent. In each replication, more attractive attorneys are significantly more likely to win individual judges' votes and the case on the merits.

The results of this analysis are relevant to multiple audiences. First, this paper is among the first studies to find evidence of plainly irrational behavior in an examination of heuristics in judicial decision-making using real-world data. Decisions

Powell Oral Argument Notes (Johnson 2009). They are available at http://users.polisci.umn.edu/~trj/oanotes93.php.

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

based on attorney attractiveness do not align with any legal, attitudinal, or strategic model of judging. These results establish convergence between observational and previous experimental data and provide multi-methodological support that judges rely on heuristics and that heuristics significantly alter their behavior. This fact complicates attempts to model their decision-making using standard approaches. These findings should serve as a strong wake-up call for judicial researchers to consider alternatives to the traditional rational choice perspective in their work.

Second, and perhaps more importantly, these results raise normative concerns that question the American ideal of equal justice under law. Attractiveness is not supposed to influence the outcome of legal disputes. The famous cliche "justice is blind" implies impartiality and objectivity in the legal system. It suggests that decisions are made based on authoritative texts, precedent, and unique case facts, rather than personal characteristics of the parties or their attorneys. However, if attractive attorneys have an inherent advantage, it raises doubts about just how fair the courtroom can be. In short, these results have serious normative implications for any person that interacts with the US justice system.

Heuristics and the judiciary

Heuristics are instinctual feelings or reactions that help humans associate new phenomena with previously observed patterns for quick decision-making. A reliance on heuristics is not necessarily problematic. The intuitions that underlie their use are responsible for keeping early humans safe from predators and modern humans away from scammers in their inbox. But they become problematic when they are used to complete complex tasks (Tversky and Kahneman 1974). When given difficult math problems or logic puzzles, the use of heuristics causes subjects to answer questions erroneously and justify their answers with faulty reasoning. This effect is not unique to "laymen"; studies have confirmed similar errors in decision-making by doctors (e.g., Nordholm 1980), teachers (e.g., Adams and Cohen 1974), and statisticians (e.g., Kahneman and Tversky 1982).

Judges, however, have traditionally been considered less susceptible to the dangers of heuristic decision-making. This view is based on two reasons given by two groups of skeptical court watchers. The first group argues that, as a result of their extensive training and experience in the legal profession, judges are immune to using heuristics in substantive decision-making. Judges may feel emotions in the course of their work, but they are able to "suppress" their intuitive thoughts and feelings and convert them into deliberately constructed rational decisions (Posner 2010, 119). To be sure, federal appellate judges are the elites of the legal community and resolve many controversial and important disputes. Most of these judges have reached the pinnacle of their profession through successful careers in litigation, academia, or both.[2] Nearly all of their education and experience in the legal profession, this group contends, discourages the use of heuristics. A judge that sacrifices deliberation for efficiency could never appreciate the legal nuances of an individual case and might be too reliant on extraneous details to adequately apply precedent, relevant statutes, or their own jurisprudence. In short, even if judges have instinctual reactions to new

[2]See the Federal Judicial Center's Biographical Directory of Article III Federal Judges, 1789-present. URL: https://www.fjc.gov/history/judges.

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

stimuli, they argue those reactions "will not flourish under the light of intense study" (Coffin 1994, 255).

But psychological research suggests that judges may actually be among the *most* susceptible to relying on heuristics. All humans have evolved to seek to minimize their cognitive stress and maximize their cognitive ease (Kahneman 2011). Repeated exposure to complex situations that require deliberate decision-making incentivizes the use of familiar heuristics to avoid exhaustion. With this in mind, consider the workload for an average judge on the United States Court of Appeals. Without the benefit of discretionary review, a judge hears an average of 418 cases per year.[3] For each case, the judge must read multiple briefs, occasionally prepare for and listen to oral argument, conference with their colleagues, and draft an opinion (or at least be prepared to read and analyze a colleague's). Even if a judge allocated every day of the year to a specific case, and was able to complete all of the required work in a 24-hour period, they would not be able to dispose of all the cases on their docket. But the sheer number of cases is not all that contributes to a judge's workload. The information on which a judge relies to reach a decision often comes from unfamiliar sources. Due to the random nature of case assignment at the Courts of Appeals, judges interact with the same attorneys less frequently than judges on other courts. Judges must make decisions both on the content and the credibility of information before them. Finally, there is a subset of cases on the judge's docket that are particularly complex, or in an area of the law the judge does not know well, with extra effort required to make an informed decision. Even with a contingent of law clerks, the totality of the work can be overwhelming. Under these conditions, it is difficult to imagine that judges can call upon the "intense study" that would allow them to avoid heuristics.

The second group of skeptics consists primarily of political scientists that study judges from a rational choice perspective. These researchers assume that judges have a set of preferences for legal policy that can be ranked with regard to the amount of utility they bring the judge, that the judge will maximize their utility when accounting for uncertainty, and that their preferences will remain stable in the short run (Boncheck and Shepsle 1997). For these researchers, even if judges use heuristics, it is as a means to make rational decisions. In the early years of judicial politics, these rational decisions equated with adherence to existing law and precedent (e.g., Corwin 1924), but in more recent decades, they are based on judges' ideological leanings (e.g., Segal and Spaeth 2002), their interactions with other actors (e.g., Epstein and Knight 1998), and their personal and professional happiness (e.g., Epstein, Landes and Posner 2013). If judges are truly rational decision-makers, every choice they make, whether it is derived from heuristics or not, is in accordance with their predefined preferences.

Studies of the consequences of heuristic decision-making in the judiciary have been few in number, but have been noteworthy in suggesting this claim is not always true.[4] Jeffrey Rachlinski, Andrew Wistrich, and Chris Guthrie have conducted numerous experiments on thousands of judges at judicial conferences and found that judges rely on heuristics in predominantly the same ways as other humans and are prone to irrational decisions as a result.[5] They find that judges are likely to use heuristics when hearing especially complex cases (Rachlinski, Guthrie and Wistrich 2007), when

[3]See the Federal Court Management Statistics Summary for the Calendar Year ending December 31, 2018. URL: https://www.uscourts.gov/statistics-reports/federal-court-management-statistics-december-2018.

[4]For a couple of prominent examples in political science, see Wedeking (2010) and Epstein, Parker and Segal (2018).

[5]See also the experiment by Spamann and Klöhn (2016).

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

deciding the guilt of minority defendants (Rachlinski et al. 2009), and when awarding damages in civil litigation (Guthrie, Rachlinski and Wistrich 2007; Rachlinski, Wistrich and Guthrie 2015). In all of these experiments, the co-authors ask judges about their preferences and then randomly give them case vignettes that manipulate only the personal characteristics of the parties and/or attorneys arguing a case, leaving all of the underlying ideological positions, case facts, and relevant legal authorities constant. The fact that the experimenters find significant differences between their treatment and control groups anyway highlights both the willingness of judges to rely on heuristics and how their use can produce decisions that are inconsistent with their preferences.

Still, while experiments are often called the "gold standard" of causal identification, relying solely on experimental results to draw conclusions leaves a researcher vulnerable to the pitfalls of any single-method research design – the results are only as believable as the method's assumptions. In all experimental settings, researchers must be aware of external validity concerns. Critics contend that experimental results "cannot by themselves establish that the effect would be similar in a different setting" (Spamann 2015, 143). Because judicial experiments must be conducted outside of the courtroom using abbreviated time scales and information packets, many judges and scholars argue they are too artificial to replicate courtroom decision-making (e.g., Maroney 2015). This paper addresses these concerns by testing for evidence of heuristic decision-making in real-world decisions.

Attractiveness as a heuristic

Psychologists believe that attractiveness is one of the most widely used heuristics in everyday decision-making. More attractive individuals enjoy advantages in many aspects of life over less attractive colleagues. In the workplace, studies have shown that better-looking people are more likely to be considered qualified applicants (Dipboye, Fromkin and Wiback 1975), be hired (Gilmore, Beehr and Love 1986), earn promotions (Hamermesh and Parker 2005), and receive higher compensation (Frieze, Olson and Russell 1991). In school, better-looking students are more likely to receive attention from their teachers (Adams and Cohen 1974), earn better grades (DeMeis and Turner 1978), and be more popular socially (Boyatzis, Baloff and Durieux 1998). In the doctors office, better-looking patients receive more attention from their doctor (Nordholm 1980) and are considered healthier (Hadjistavropoulos, Ross and Von Baeyer 1990), while better-looking doctors are trusted more by their patients (Brase 2004). Research even suggests that attractive people live less stressful, happier lives (Gupta, Etcoff and Jaeger 2016).

While the study of attractiveness has entered political science, much of this work has focused on election outcomes. For example, researchers have found that candidates judged to have more "competent" faces are more likely to have electoral success (Todorov et al. 2005) and that these same candidates are also more likely to succeed in markets with more television advertising in advance of an election (Lenz and Lawson 2011). Stockemer and Praino (2017) find that the effect of appearance is most pronounced in an electoral system that incentivizes straightforward decision-making (e.g., first past the post) and mitigated in more complex systems (e.g., single transferable voting).[6]

[6]Despite this work in neighboring areas of study, there have been very few studies of the effect of appearance on elites in the legal world. Stewart (1980) examines the sentences laid down in 73 cases by

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

All of the advantages detailed here are born out of the "beauty is good" principle. This principle applies the halo effect, the tendency for positive impressions of a person in one area to positively influence one's opinion or feelings in other areas (Thorndike 1920), to the physical quality of attractiveness. Psychologists Karen Dion, Ellen Berscheid, and Elaine Walster were the first to theorize that humans use one another's appearance as a shortcut to make a judgement on qualities that are difficult to observe. In an experimental setting, they find support for their theory; attractive individuals are assumed to be more socially desirable and more successful simply because they are attractive (Dion, Berscheid and Walster 1972). Benoît Monin extended this principle to include "beauty is familiar." Monin argues that attractive people are not just assumed to be successful by others; they engender positive feelings because they are intuitively viewed by others as both familiar and trustworthy (Monin 2003).

While there has been no empirical study of the subject, there is considerable anecdotal evidence to suggest that federal judges evaluate the attractiveness of the people they interact with just like everyone else. The most well-known evidence to political scientists comes from the papers of the aforementioned Justice Harry Blackmun. For almost all of the oral arguments in which he participated, Blackmun made a note on the appearance of each attorney appearing at the lectern. They were honest evaluations of the attorney's appearance that ranged from positive ("very nice looking"), to negative ("fat, little oily"), to neutral ("in red and ribbon").[7] Even apart from Justice Blackmun, appearance has been specifically mentioned as crucial to establishing credibility by multiple judges in advice articles to law students and practicing lawyers (e.g., Aldisert 2010; Quinn 2012). Further, suggestions for an attorney's appearance are explicitly mentioned in the guidelines for oral argument at both the Supreme Court (Stern et al. 2002, 687) and Circuit Court level (Martineau et al. 2005, 872).

This anecdotal evidence combined with the empirical results in neighboring fields of study reviewed herein leads to strong expectations for the effect of attorney attractiveness on judicial decision-making. I suspect that judges rely on an attorney's attractiveness to make decisions simply because judges are human. A judge is required to make many decisions based on information that she receives from attorneys. Many of these attorneys are unfamiliar and require the judge to evaluate their credibility. Determining a stranger's credibility is difficult and time consuming for anyone, but particularly for a busy judge. The Rachlinski et al. experimental research has shown that a judge will use an easily recognizable quality like gender, race, or income as a

judges in Pennsylvania, finding that judges sentenced attractive defendants to shorter sentences than unattractive defendants. Downs and Lyons (1991) find that judges were more lenient on attractive defendants when levying fines and setting bail. Lastly, Zebrowitz and McDonald (1991) find that attractive and "baby-faced" litigants were more likely to succeed in a sample of small claims courts. While these studies are a promising start, they rely on small-N samples of state court judges. They also focus solely on the appearance of defendants without considering the role of the attorney arguing the case (in some cases, there is no attorney is present in the courtroom). To date, there has been no study of the effect of attractiveness at the highest levels of the judiciary, the federal appellate courts.

[7]As with the prior examples, these notes from *M/S Bremen* v. *Zapata Off-Shore Co.* 1972, *Craig* v. *Boren* 1976, and *Hughes Tool Co.* v. *TWA* 1972, respectively, can be viewed at the The Digital Archives of Justices Blackmun and Powell Oral Argument Notes (Johnson 2009).

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

substitute for harder-to-identify qualities relevant to legal decision-making. I expect attorney attractiveness to have a similar effect. When appearing before the judge at oral argument, an attorney's attractiveness will serve as a signal of the desirable qualities the judge is seeking in their argument. These qualities may include coherent legal reasoning, proper treatment of precedent, or ideological agreement with the judge, but none of those things need actually be present. In a close case, the attribution of the desirable qualities to an attorney's argument is enough to be the difference between winning a judge's vote and losing it.

This gives rise to the following hypothesis:

H1: When compared to an otherwise similar opponent, a more attractive attorney will be more likely to win a judge's vote than a less attractive attorney.

Still, judges at the appellate levels of the US judicial hierarchy do not decide cases on their own. Most cases at the US Courts of Appeals are heard by three-judge panels, while en banc and US Supreme Court cases are heard by even larger courts. To fully understand the policy consequences of using attractiveness as a heuristic, it is necessary to consider the case-level effect. Because the attractiveness advantage is extended by humans regardless of other characteristics such as their race or gender (e.g., Monin 2003), I expect the success of attractive attorneys to persist across an entire panel of judges.

This gives rise to an additional hypothesis:

H2: When compared to an otherwise similar opponent, a more attractive attorney will be more likely to win a case than a less attractive attorney.

Causal inference on the effect of attractiveness

To best identify causal effects, many empirical researchers rely on the "but for" method of statistical analysis. That method, applied to this paper, asks "but for the fact that an attorney is considered attractive, would the attorney still win a judge's vote?" To answer this question, I use a matching algorithm to create a dataset of like attorneys, balanced on characteristics that may predict their probability of success, with the only point of difference being their attractiveness. In the language of causal inference, an attorney's attractiveness is the "treatment," while the balancing covariates are characteristics that include the ideological position of an attorney's argument, the attorney's previous experience at oral argument, and shared demographic characteristics with the judges casting votes. The quantity of interest is the sample average treatment effect on the treated (ATT), the increase in the likelihood of receiving a judge's vote due to attractiveness.

The best environment for causal identification is one in which the researcher has complete control over all of the variables relevant to the outcome variables. Ideally, I would observe the outcome after an attractive attorney argued a case before a judge, go back in time and have the case re-argued by a similarly qualified but less attractive attorney, observe the outcome again, and compare the differences. Of course, this is not feasible, and a replication of this situation in a laboratory setting is not favorable given the need for externally valid data in this analysis. To test for a causal relationship, I use techniques that allow observational data to mimic the advantageous qualities of an experiment and require minimal assumptions.

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

Dataset, treatment, and covariates

To test my hypothesis, I compiled a dataset of all orally argued cases at the US Courts of Appeals from 2017 to 2019 in which the US government was a party. I rely on the Courts of Appeals because the panels of judges assigned to hear cases are assigned without any input from the parties or the judges. Attorneys are unable to select judges who may be more or less favorable to their case, and judges are unable to select cases argued by attorneys they more or less prefer. In addition, oral argument at the Courts of Appeals provides an interesting environment to observe heuristic behavior. Oral argument is not mandated in any of the federal circuits on any issue area and usually occurs only in "novel areas of the law" (Wasby 1981, 56). In the 2017–2019 time period less than 20 percent of cases were selected for oral argument. The rarity of oral argument and the randomness of assignment also lead to more unfamiliar attorneys arguing these novel cases.[8] In these conditions, where they are forced to make quick decisions on credibility in a complex case, judges are likely to lean on heuristics.

The analysis is restricted to the three-year period from 2017 to 2019 for reasons of internal validity. To conceptualize attractiveness, I rely, in part, on images of attorneys obtained from their professional websites. As I conducted my analysis in 2020, beginning the sample in 2017 ensures that (1) the vast majority of attorneys in the data had images available on their websites and (2) the images are a fair depiction of the attorneys' appearance on the day they were in court. Prior to 2017, it becomes increasingly difficult to find images of attorneys that accurately portray their appearance on the day of oral argument or are of equal quality to the images of attorneys from 2017 forward.

I rely on the US government as a common opponent to facilitate a comparison between attorneys with varying attractiveness ratings. I assume that the attorney representing the US government receives an advantage with the judges regardless of factors that may otherwise connote attorney quality (e.g., experience or law school pedigree). This assumption is supported by existing Courts of Appeals research that shows the US government has a consistent advantage over its opponents (e.g., Cross 2007). This approach allows me to avoid the complex interpretation required when analyzing two attorneys that have different attractiveness ratings and different values in other covariates predictive of success at the Courts of Appeals. Rather than comparing attractive and unattractive attorneys when they face one another, I compared their performance against a comparable opponent, providing a clear interpretation of the ATT.[9]

The dataset contains cases from the eleven numbered circuit courts and the D.C. Circuit. In total, it contains 1,067 unique cases, 930 unique opposition attorneys representing clients litigating against the US government, and 466 unique US attorneys. The judge-level outcome variable, **Opposition Attorney Vote**, is a binary variable that records a judge's vote in favor of the opposition attorney. The dataset contains 3,290 judge votes. Of those votes, 1,448, or 44 percent, are in favor of the opposition attorney. The case-level outcome variable, **Opposition Party Win**, is a binary variable that records a Courts of Appeals panel's decision in favor of the

[8]Nearly 25 percent of the attorneys in this analysis were arguing their first case at the Courts of Appeals.

[9]In addition, Nelson and Epstein (2021) suggest that this approach may generalize the findings of this project to other courts because of the persistent presence (e.g., Weinshall and Epstein 2020), and success (e.g., Alarie and Green 2017), of federal parties in courts throughout the world.

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

Table 1. Descriptive Statistics for Variables of Interest

	Mean	Std. dev.	Respondent	Image source	Circuit(s)
Attractiveness Rating (Images)	6.624	0.899	Humans	Attorney websites	All
Attractiveness Rating (Computer)	7.719	0.611	Computer	Attorney websites	All
Attractiveness Rating (9th)	5.243	1.133	Humans	Courtroom videos	9th
Opp-US Attractiveness Difference	–0.622	1.336	Humans	Courtroom videos	9th

opposition party. Of the 1,067 unique cases in the dataset, 449 or 42 percent, are decided favor of the opposition party.[10]

I collect four different continuous measures of attractiveness that allow me to account for potential differences in what humans and computers deem attractive, how individuals are rated in professional images compared to courtroom videos, and whether the advantages of attractiveness are observed individually or relative to other individuals. Table 1 summarizes the characteristics of the different treatment variables.

The first measure is derived from survey respondents' ratings of physical attractiveness based on photos obtained from attorneys' professional websites. The photos are standardized in terms of size, color, and background relative to other photos in the dataset.[11] Figure 1 provides an example of two photos used in the survey. The respondents were recruited through Amazon's Mechanical Turk ("MTurk") platform and gave evaluations of an attorney's attractiveness on a 1 (not attractive) to 10 (very attractive) scale.[12] On average, each attorney was given a rating by ten different respondents. The ratings were then averaged to create the continuous treatment variable **Attractiveness Rating (Images)**. In total, 1,185 unique attorneys received an attractiveness rating.

The second measure uses the professional images in combination with existing machine learning tools designed to evaluate facial attractiveness. Specifically, I made

Figure 1. Examples of Professional Attorney Images.

[10]The precise distribution of the outcome variable differs slightly in the different analyses presented in this paper. See Table A.1 of the online appendix.

[11]This standardization follows the advice given by Todorov et al. (2005) among others.

[12]This research adheres to all Principles and Guidance for Human Subjects Research approved by the APSA Council in Spring 2020. See the online appendix for a detailed discussion.

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

Figure 2. Examples of Still Shots From Ninth Circuit Videos.

use of the Face++ Beauty Score algorithm that was created by the technology company Megvii. Face++ evaluates qualities like facial symmetry, hairstyle, and teeth alignment to assign each image a score from 1 (not attractive) to 10 (very attractive).[13] The ratings were used to create the continuous variable **Attractiveness Rating (Computer)**. Megvii imposes additional requirements for accurate ratings such that individuals with large glasses or subjects looking away from the camera cannot be included. Thus, some attorneys used in the full analysis were excluded, leaving 743 unique attorneys that received ratings from the algorithm.

A third measure addresses the possibility that an attorney had a photograph several years old on their website or that they were having a particularly bad hair day on the day of argument. If the image viewed by the survey respondents and the appearance viewed by the judge are not similar, then the attractiveness ratings could be biased in an unpredictable direction.

I took advantage of the video recordings of oral argument made publicly available by the Ninth Circuit Court of Appeals. To use these videos, I created an additional dataset of all orally argued cases at the Ninth Circuit in which the US government was a party between 2017 and March 2020. The Ninth Circuit dataset contains 454 unique cases, 1,522 judge votes, 417 unique opposition attorneys, and 252 unique US attorneys.

The attractiveness ratings for attorneys arguing in the Ninth Circuit were obtained identically to the imaged-based ratings with one exception. Survey respondents rated an attorney's attractiveness based on a ten-second, silent, video clip collected from the Ninth Circuit's YouTube page.[14] This guarantees that the respondents rated an attorney's appearance as it was on the day the attorney was in court. Figure 2 provides an example of still shots from the video recordings. The attorney's ratings were averaged to create the continuous **Attractiveness Rating (9th)**. Each attorney received a new attractiveness rating for every case that they argued in the Ninth Circuit for a total of 633 unique ratings.

A fourth and final measure allows for direct comparison between competing attorneys at oral argument. While I assume that all US attorneys are comparable opponents by virtue of their office, it may be possible that judges evaluate their attractiveness similarly to the opposition attorneys and side with the more attractive US attorneys more frequently. To account for this possibility, I had survey

[13]A free version of the algorithm can be found at https://www.faceplusplus.com/beauty/.

[14]The Ninth Circuit's YouTube page can be found at https://www.youtube.com/c/9thCircuit/videos.

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

respondents rate the attractiveness of every attorney's US government opponent in the Ninth Circuit sample.[15] The treatment variable, **Opp-US Attractiveness Difference**, is the difference between the opposition attorney's attractiveness score and the US attorney's score. Positive values indicate an advantage for the opposition attorney, while negative values indicate an advantage for the US attorney. As before, each attorney, opposition and US, received a new attractiveness rating for every case they argued before the Ninth Circuit. In total, 633 pairs of attorneys were rated for this analysis.

While distinct, the four measures are all positively correlated at a statistically significant level and are distributed similarly, suggesting there is an underlying attractiveness the persists in every measure.[16] Importantly, in order to preserve a sufficient sample size, only Attractiveness Rating (Images) is used in the case-level analysis.

The covariates used in the matching analysis and the models of opposition attorney success are familiar to the Courts of Appeals decision-making literature.[17] They include the following: **Experienced Attorney**, indicating if the attorney has argued a case at the Courts of Appeals before; **Amicus Advantage**, measuring net amicus curiae support; **Elite Law School Graduate**, measuring if the opposition attorney attended an elite law school; **Opposition Party Resources**, categorically ranking the resources available to the non-US party; and **En banc**, indicating if the case was heard en banc. To account for any gender or racial effects in the rating of the attorneys by survey respondents, I also include **Shared Gender with Respondent** and **Shared Race with Respondent** in the matching analysis. In the judge-level analysis, **Unfavorable Judge** measures if the judge is ideologically pre-disposed to vote against the opposition attorney; along with **Shared Gender with Judge**. In the case-level analysis, **Unfavorable Panel** measures if the majority of a panel is ideologically pre-disposed to decide against the opposition party; **Shared Gender with Panel** captures if the attorney shares a gender with the majority of the panel; and **Shared Race with Panel** captures if the attorney shares a race with the majority of the panel.

Empirical strategy

A straightforward logistic regression model of the effect of attractiveness on attorney success with all control variables shows that more attractive attorneys have a positive and statistically significant advantage over less attractive attorneys across all four measures of attractiveness.[18]

However, this information is not particularly helpful in learning about the causal effect of attractiveness. Because attorneys are not randomly assigned the treatment variable, natural imbalance exists in the data. As attractive attorneys, like attractive non-attorneys, receive advantages in schooling and employment, the set of covariates used in a basic regression model are not likely to be equal across the treatment and

[15]I used the Ninth Circuit dataset for this analysis to ensure the quality of images used to create the unique attractiveness scores were comparable. Many US attorneys do not have professional websites and, accordingly, do not have professional headshots like the opposition attorneys in private practice.

[16]See Table A.2 and Figure A.1 in the appendix for further detail.

[17]Complete descriptions of the covariates are included along with descriptive statistics in the online appendix.

[18]The full results of the basic model are displayed in Table A.6 of the appendix.

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

control groups (Dion, Berscheid and Walster 1972). Indeed, in an unmatched full dataset, 44 percent of attorneys with above-average computer-based attractiveness scores attended elite law schools compared to 29 percent of below-average attorneys.

While this imbalance complicates the ability to estimate the ATT, there are remedies that researchers can impose on their data. Many researchers, for example, dichotomize their continuous treatment variables to meet the requirements of matching algorithms. But this can be a problematic transformation for two reasons. First, much of the variation in the treatment variable is lost. It has been shown that dichotomizing a variable at the median reduces statistical power by the same amount as discarding a third of the data (MacCallum et al. 2002). Second, the transformation is reliant on a researcher-defined cutpoint that has been shown to dramatically affect the results of any analysis performed (Fong, Hazlett and Imai 2018).

With these concerns in mind, I use a generalized version of covariate balancing propensity score matching (CBPSM) that allows for continuous treatment variables to balance my data. The intuition of all propensity score matching is to calculate the probability that an observation will be assigned the treatment based on other observed covariates (Rosenbaum and Rubin 1983). This probability is called a propensity score. Propensity score matching matches treated units with untreated units based on a similar propensity score. (Imai and Ratkovic (2014) introduce CBPSM to optimize the balance of the covariates' values across treated and control units after matching in addition to the probability of treatment assignment. Fong, Hazlett and Imai (2018) extend CBPSM to allow for continuous treatments by estimating propensity scores such that covariate balance is optimized automatically (as opposed to having the researcher manually check balance statistics with every new value of a treatment variable). The Fong, Hazlett, and Imai version of CBPSM results in better covariate balance than standard propensity score matching and reduces sensitivity to misspecification, all while allowing researchers to take advantage of the insights available from the increased variation in their treatment variable. Figure 3 displays the results of CBPSM on unique datasets used for the four continuous measures of attractiveness at the judge-level and the image-based measure at the case-level. The results are shown in the form of correlations between the treatment variable and the covariates. Correlations farther from zero indicate greater imbalance between the treatment and control groups. For a perfectly balanced covariate, the correlation is equal to zero. For every conceptualization of the treatment variable, each covariate's balance is improved after CBPSM.

With balanced datasets, the analyses are straightforward. Because the dependent variable is dichotomous, I fit a series of logistic regression models with a set of pre-treatment covariates for each measure of the treatment. I rely on this strategy to assume that the outcome variable is independent of treatment status given the covariates (e.g., Ho et al. 2007).

Results

Judge-level effects

I fit four logistic regression models on the balanced datasets for each of the four measures of attractiveness. Table 2 displays the coefficients of the logistic regression models and the odds ratios after CBPSM matching. The results are consistent with the expectation of H1. The raw coefficients are useful for determining the direction

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

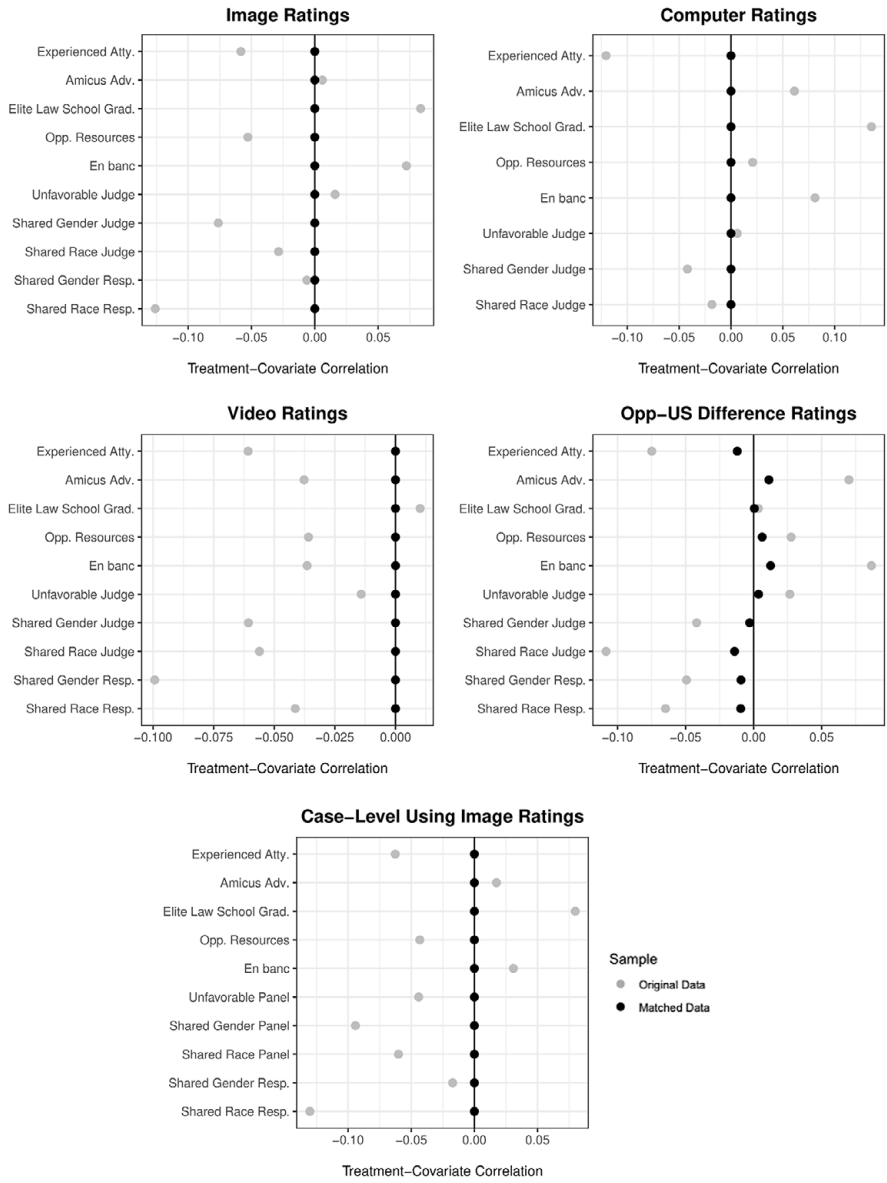

Figure 3. Balance Statistics After CBPSM.
Note: This figure displays the correlations between the values of the covariates and the judge-level treatment variables. Black dots indicate the correlation between the variables in the matched datasets, while gray dots indicate the correlation in the unmatched datasets. Correlations between a treatment and the covariates close to 0 indicate greater balance in the data. See Tables A.7–A.11 in the online appendix for tabular versions of this figure.

and the statistical significance of an effect. All four measures of attractiveness are positive and statistically significant, indicating attorneys with higher attractiveness ratings (or greater advantages over their US opponents) are more likely to receive a judge's vote than their less attractive counterparts. The odds ratios, found by

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

Table 2. The Effect of Attractiveness on Receiving a Judge's Vote

	Dependent variable:				
	Opposition Attorney Vote				Odds ratio
Attractiveness Rating (Images)	0.107*				1.113
	(0.040)				
Attractiveness Rating (Computer)		0.255*			1.291
		(0.067)			
Attractiveness Rating (9th)			0.181*		1.198
			(0.050)		
Opp-US Attractiveness Difference				0.087*	1.091
				(0.041)	
Control variables	✓	✓	✓	✓	
Constant	–1.459*	–2.757*	–1.159*	–0.223	
	(0.301)	(0.540)	(0.381)	(0.275)	
Observations	3,290	2,646	1,522	1,522	

Note: logistic regression models after CBPSM; one-sided t-tests; *p <0.05;

exponentiating the raw coefficients, can be used for an interpretation of the effect's magnitude. They are interpreted as follows: a one-unit increase in an attorney's image-based attractiveness score makes the attorney 1.113 times more likely to receive a judge's vote. While at first glance this effect size seems small, the continuous form of this treatment variable affects the interpretation. For *each* one-unit increase in an attorney's attractiveness score, they become 1.113 times more likely to receive a vote. This means an attorney with a score of 8 has a small advantage over an attorney with a score of 7, but a much greater advantage over an attorney with a score of 4. The size of the effect is similar with respect to the other attractiveness measures.

The interpretations of the treatment effects are more intuitive when they are translated into the predicted probability plots displayed in Figure 4. In each plot, as an attorney's attractiveness rating increases, so does the probability of receiving a judge's vote when the other covariates are held constant. For an attorney with an image-based attractiveness score of 3 (the minimum score in the data), the probability of winning a judge's vote is 0.351. This probability increases to 0.402 with a score of 5, 0.454 with a score of 7, and 0.507 with a score of 9 (the maximum in the data). This equates to approximately a 5 percent increase in the likelihood of success per a two-unit increase in attractiveness scores. This compares to an 11 percent rate for the computer-based scores, 8 percent for the video-based scores, and 5 percent for a two-unit increase in the opposition and US attorney difference.

See Table A.12 in the online appendix for full results.

Case-level effects

For the case-level effect of attractiveness, I fit a logistic regression model with **Opposition Party Win** as the dependent variable, the image-based attractiveness score as the treatment, and case-level covariates balanced after CBPSM. Table 3 displays the attractiveness coefficient from the logit model and the odds ratio. The results are consistent with the expectation of H2. Attractive attorneys are significantly more likely to win their cases than their less attractive counterparts. The odds ratio reveals that a one-unit increase in an attorney's image-based attractiveness score

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

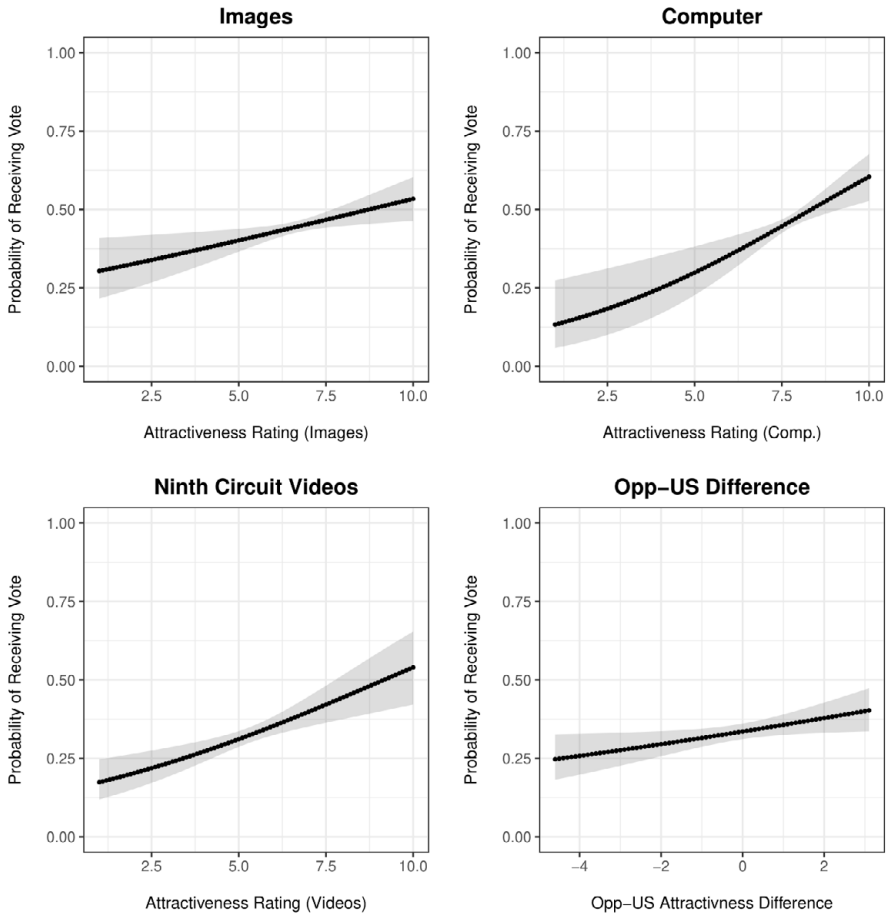

Figure 4. Predicted Probability of Receiving a Judge's Vote.
Note: Probabilities are calculated after CBPSM. The shaded areas represent 95 percent confidence intervals. All ratings are derived from survey respondent's evaluation of images of attorneys on a 1 (low attractiveness) to 10 (high attractiveness) scale. All of the control variables used in the analyses are held at their mean values to create this figure. See Table A.13 in the appendix for the table on which this figure is based.

makes the attorney 1.135 times more likely to win a case. This is a similar size to the judge-level effect.

See Table A.14 of the online appendix for full results.

Figure 5 displays the predicted probability of winning a case after CBPSM. When all of the other covariates in the model are held constant, an attorney with an image-based attractiveness score of 3 has probability 0.312 of winning a case. This probability increases to 0.368 with a score of 5, 0.429 with a score of 7, and 0.492 with a score of 9. This equates to a roughly 6 percent increase in the likelihood of success with every two-unit increase in attractiveness scores.

Subsequent investigation of the predicted probability results reveals that in 213 cases, or approximately 20 percent of the cases in this sample, an attorney's above-average treatment score was the difference between winning and losing a case.

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

Table 3. The Effect of Attractiveness on Winning a Case

	Dependent variable:	
	Opposition Party Win	Odds ratio
Attractiveness Rating (Images)	0.126*	1.135
	(0.071)	
Control variables	✓	
Constant	−1.820*	
	(0.530)	
Observations	1,067	

Note: logistic regression model after CBPSM; one-sided t-test; *p <0.05;

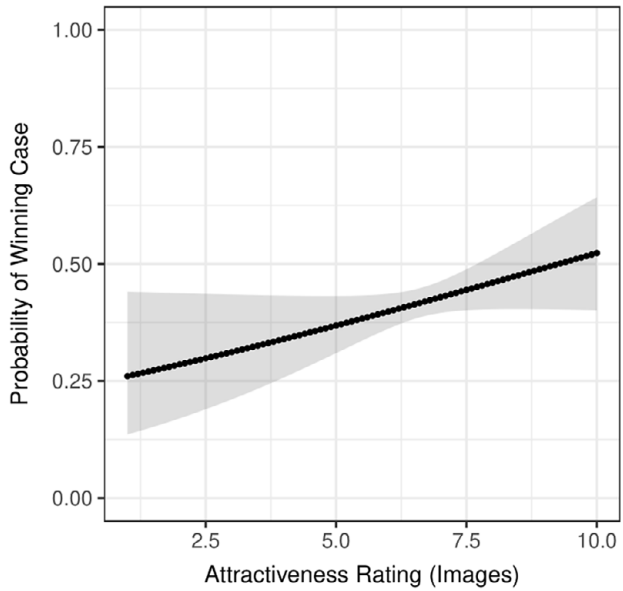

Figure 5. Predicted Probability of Winning a Case.
Note: Probability is calculated after CBPSM. The shaded area represents a 95 percent confidence interval. All of the control variables used in the case-level analysis are held at their mean values to create this figure. See Table A.15 in the appendix for the table on which this figure is based.

When the coding of the treatment variable is reduced by one standard deviation in these cases, the predicted result of the cases diverges from the real-world result.

Robustness checks

In addition to comparing various conceptualizations of attorney attractiveness, I also perform five robustness checks: replicating the analysis with Coarsened Exact Matching (CEM), introducing sampling weights to have the sample of survey respondents approximate the sample of judges in the full dataset, and incorporating additional covariates and fixed effects in the analysis.

CBPSM allows me to take advantage of the continuous nature of the treatment variables, but there are many other methods available to researchers who decide to

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

dichotomize their treatment variable that have certain advantages. To use one of these methods, I dichotimize the treatment variables to create "high attractiveness" indicators. A high attractiveness score is a score that is more than one standard deviation above the average attractiveness score of all the attorneys in a given sample. I also create a variable of attractiveness advantage to indicate opposition attorneys that have higher attractiveness scores than their US government opponents. With binary treatment variables, Iacus, King and Porro (2019) show that causal effects can be estimated from any data assumed to be generated by a stratified random sampling framework (also known as block randomization) as opposed to simple random sampling (used in propensity score matching). For this robustness check, I adopt this axiom and balanced my dataset absent assumptions about the covariates' probability distributions using a post-stratification based matching approach. While any of the multiple Monotonic Imbalance Bounding (MIB) methods are suitable to accomplish this task (Iacus, King and Porro 2011), I replicated my analysis using CEM. The substantive results remain the same. The complete results for this analysis are included in Section 5.1 of the appendix.

To account for differences between the demographic characteristics of serving US Courts of Appeals judges and the survey respondents responsible for assigning attractiveness scores, I assign weights to attractiveness scores based on a respondent's self-reported gender and race. In this weighted sample, the amount of influence held by each gender-race combination is identical to the influence held by that combination in the sample of US Courts of Appeals judges included in the analysis. The relevant continuous treatment variable is an image-based weighted attractiveness rating that is an average of the weighted scores unique to each attorney. The mean is 6.501, slightly less than the unweighted rating mean of 6.624. After CBPSM is applied to the dataset and a logistic regression model is fit, the substantive results hold. The complete results for this analysis are included in Section 5.2 of the appendix.

All matching algorithms are only as helpful as the covariates provided by the researcher. Matching algorithms may mimic experiments, but they cannot resolve omitted variable bias entirely because the treatment is not randomly assigned. All researchers must weigh the trade-offs between sample size and a balanced dataset when deciding how they will use matching in their analysis. While I have presented one specification in this project, there are others that, in particular, allow for more nuanced attorney experience and ideological variables. To account for these alternatives, I collected a continuous version of **Unfavorable Judge** that accounts for the ideological distance between the judge and the opposition party's position based on their position relative to the lower court judge. In addition, I include binary variables on the opposition attorney's appellant status, the opposition attorney's prior service as a Courts of Appeals law clerk, and the opposition attorney's prior service as a Supreme Court law clerk. Lastly, to account for unique effects that may be based on the rules, customs, or traditions of any particular circuit, I include circuit fixed effects for matching and empirical analysis. After adding these variables to the full dataset, I performed CBPSM with the image-based attractiveness treatment and fit a logistic regression model to the balanced data. The results are unchanged; more attractive attorneys remain more likely to win a judge's vote and a case. The complete results for this analysis are included in Section 5.3 of the appendix.

Finally, to account for the different perceptions of attractiveness that can be influenced by a combination of cultural, societal, and individual factors related to gender and race that are difficult to measure with a single set of control variables, I

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

also sought other ways to account for the opposition attorneys' demographic characteristics. In the main analysis, I account for the theoretical expectations that a shared (or not shared) race and gender will have on attractiveness evaluations and success in court (see Meissner and Brigham 2001, for a review of many of these studies). While this approach is useful, an alternative approach would directly control for the role of disparate races and genders in the analysis. As a robustness check, I added opposition attorney gender and minority racial status as two additional matching covariates and control variables. Even when attorneys are matched on their race and gender, the more attractive attorney is more likely to have success in court. A further examination of possible conditional effects reveals interesting, albeit preliminary, results. Li (2015) among others suggests that any attractiveness advantage is likely to be rewarded at a different rate to male and female attorneys. In this sample, although the individual attractiveness advantage given to attorneys based on their gender is mixed, there is evidence of a three-way conditional effect of greater advantages being extended to female attorneys with a high attractiveness score when arguing before male judges. These initial findings suggest avenues for future research to more fully understand the nuances of this advantage. The complete results for this analysis are included in Section 5.4 of the appendix.

Discussion

Former Court of Appeals Judge Howard T. Markey once said that "all people have values of some sort, and judges are people. [A judge] without values would be a robot, an automation" (O'Brien 1981, 285). While it is difficult to draw conclusive evidence from one observational study, the results presented here suggest that Markey was right; judges are people, and they make decisions in accordance with at least one of the values all people seem to share – beauty is good. Some judges may disagree, but it is not too surprising to find evidence that judges use heuristics in their decision-making. After all, political scientists have found evidence of their use previously without always knowing what they were looking at. Consider how political scientists like Segal and Spaeth (2002) have shown the effect of ideology on judicial decision-making and how subsequent studies have shown the shortcuts judges use to identify a party's ideology. Judges use heuristics when they make decisions based on the positions of partisan amicus curiae (e.g., Collins 2008), the position of the US government (e.g., Black and Owens 2012), or the position of their colleagues on the bench (e.g., Kastellec 2020). In these examples, judges use heuristics to arrive at undoubtedly rational decisions.

It is more surprising to learn that heuristics can also result in irrational decision-making. In this analysis, I have shown that attorney attractiveness significantly increases the chances of attorney success even when controlling for ideology, attorney quality, and other relevant case characteristics. Whether one subscribes to a legal, attitudinal, or strategic theory of judicial behavior, voting on the basis of attractiveness does not constitute rational decision-making. The fact that my substantive findings converge with the results of the Rachlinski et al. experiments provides evidence that rational choice models cannot account for all aspects of judicial behavior. Moving forward, I suggest that judicial behavior scholars need to place greater emphasis on the parties involved in the case. As Epstein, Parker and Segal (2018, 241) surmise, this has not been "standard operating procedure" in judicial

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

politics where case-level variables have always been given priority. But as multi-methodological evidence now suggests, and the theories of psychology and behavioral economics support, the attorneys and the parties they represent are not just vehicles delivering an ideological message. They are humans whom judges interact with and evaluate. After all, the study of judging, like all social science, is ultimately the study of human interaction.

In addition, this specific case of irrational behavior is of great normative importance to the courts. The American justice system's legitimacy is maintained on the premise that cases with the same legal questions, precedents, statutes, and facts are decided in the same way. The ideal system, particularly at the trial court level, is essentially objective and therefore impartial. But this cannot be possible when attractive attorneys have an advantage in the courtroom when the characteristics of individual cases and attorneys are held constant. Attractiveness is, at least to an extent, subjective (Weeden and Sabini 2007), and relying on it opens opportunities for unfairness. Future research should examine the downstream effect of this finding. For example, are certain parties systematically able to obtain more attractive attorneys, and does this contribute to inequality in the justice system?

Lastly, there is no reason to expect the troubling consequences of using heuristics are unique to judges and the researchers who study them. Lawmakers in both the US Congress and foreign legislatures frequently operate under strong time and information constraints and, consequently, rely on heuristics when casting votes (e.g., Kropp 2010). Other work has found that the use of heuristics can explain important foreign policy decisions by American presidents and foreign leaders (e.g., Mintz 2004). These examples all lead to rational behavior by the political actors, but what of the instances where heuristics lead to inexplicable behavior? Any area of political science that relies exclusively on rational choice perspectives should consider the role of extra-political factors in elites' decision-making and how to incorporate them into their modeling strategies. All of these questions bring their own theoretical and empirical challenges, but they also bring the prospect of a more complete understanding of political decision-making.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/jlc.2024.2.

References

Adams, Gerald R., and Allan S. Cohen. 1974. "Children's Physical and Interpersonal Characteristics That Effect Student-Teacher Interactions." *Journal of Experimental Education* 43(1): 1–5.

Alarie, Benjamin, and Andrew J. Green. 2017. *Commitment and Cooperation on High Courts: A Cross-Country Examination of Institutional Constraints on Judges*. Oxford: Oxford University Press.

Aldisert, Ruggero J. 2010. *Winning on Appeal: Better Briefs and Oral Argument*. Boulder, CO: National Institute for Trial Advocacy.

Black, Ryan C., and Ryan J. Owens. 2012. *The Solicitor General and the United States Supreme Court: Executive Branch Influence and Judicial Decisions*. Cambridge: Cambridge University Press.

Boncheck, Mark S., and Kenneth A. Shepsle. 1997. *Analyzing Politics: Rationality, Behavior, and Institutions*. New York: W.W. Norton.

Boyatzis, Chris J., Peggy Baloff, and Cheri Durieux. 1998. "Effects of Perceived Attractiveness and Academic Success on Early Adolescent Peer Popularity." *Journal of Genetic Psychology* 159(3): 337–344.

Brase, Gary L., and Jillian Richmond. 2004. "The White–Coat Effect: Physician Attire and Perceived Authority, Friendliness, and Attractiveness." *Journal of Applied Social Psychology* 34(12): 2469–2481.

Coffin, Frank. 1994. *Appellate Practice and Procedure*. New York: W.W. Norton.
Collins, Paul M. 2008. *Friends of the Court: Interest Groups and Judicial Decision Making*. New York: Oxford University Press.
Corwin, Edward S. 1924. "Constitutional Law in 1922-1923: The Constitutional Decisions of the Supreme Court of the United States in the October Term, 1922." *American Political Science Review* 18(1): 49–78.
Cross, Frank B. 2007. *Decision Making in the US Courts of Appeals*. Stanford, CA: Stanford University Press.
DeMeis, Debra Kanai, and Ralph R. Turner. 1978. "Effects of Students' Race, Physical Attractiveness, and Dialect on Teachers' Evaluations." *Contemporary Educational Psychology* 3(1): 77–86.
Dion, Karen, Ellen Berscheid, and Elaine Walster. 1972. "What is Beautiful is Good." *Journal of Personality and Social Psychology* 24(3): 285.
Dipboye, Robert L., Howard L. Fromkin, and Kent Wiback. 1975. "Relative Importance of Applicant Sex, Attractiveness, and Scholastic Standing in Evaluation of Job Applicant Resumes." *Journal of Applied Psychology* 60(1): 39.
Downs, A. Chris, and Phillip M. Lyons. 1991. "Natural Observations of the Links between Attractiveness and Initial Legal Judgments." *Personality and Social Psychology Bulletin* 17(5): 541–547.
Epstein, Lee, Christopher M. Parker, and Jeffrey A. Segal. 2018. "Do Justices Defend the Speech They Hate? An Analysis of In-Group Bias on the US Supreme Court." *Journal of Law and Courts* 6(2): 237–262.
Epstein, Lee, and Jack Knight. 1998. *The Choices Justices Make*. Washington, DC: CQ Press.
Epstein, Lee, William M. Landes, and Richard A. Posner. 2013. *The Behavior of Federal Judges: A Theoretical and Empirical Study of Rational Choice*. Cambridge, MA: Harvard University Press.
Fong, Christian, Chad Hazlett, and Kosuke Imai. 2018. "Covariate Balancing Propensity Score for a Continuous Treatment: Application to the Efficacy of Political Advertisements." *Annals of Applied Statistics* 12(1): 156–177.
Frieze, Irene Hanson, Josephine E. Olson, and June Russell. 1991. "Attractiveness and Income for Men and Women in Management 1." *Journal of Applied Social Psychology* 21(13): 1039–1057.
Gilmore, David C., Terry A. Beehr, and Kevin G. Love. 1986. "Effects of Applicant Sex, Applicant Physical Attractiveness, Type of Rater and Type of Job on Interview Decisions." *Journal of Occupational Psychology* 59(2): 103–109.
Gupta, Nabanita Datta, Nancy L. Etcoff, and Mads M. Jaeger. 2016. "Beauty in Mind: The Effects of Physical Attractiveness on Psychological Well-Being and Distress." *Journal of Happiness Studies* 17(3): 1313–1325.
Guthrie, Chris, Jeffrey J. Rachlinski, and Andrew J. Wistrich. 2007. "Blinking on the Bench: How Judges Decide Cases." *Cornell Law Review* 93: 1.
Hadjistavropoulos, Heather D., Michael A. Ross, and Carl L. Von Baeyer. 1990. "Are Physicians' Ratings of Pain Affected by Patients' Physical Attractiveness?" *Social Science & Medicine* 31(1): 69–72.
Hamermesh, Daniel S., and Amy Parker. 2005. "Beauty in the Classroom: Instructors' Pulchritude and Putative Pedagogical Productivity." *Economics of Education Review* 24(4): 369–376.
Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15 (3): 199–236.
Iacus, Stefano M., Gary King, and Giuseppe Porro. 2011. "Multivariate Matching Methods That Are Monotonic Imbalance Bounding." *Journal of the American Statistical Association* 106(493): 345–361.
Iacus, Stefano M., Gary King, and Giuseppe Porro. 2019. "A Theory of Statistical Inference for Matching Methods in Causal Research." *Political Analysis* 27(1): 46–68.
Imai, Kosuke, and Marc Ratkovic. 2014. "Covariate Balancing Propensity Score." *Journal of the Royal Statistical Society: Series B*: 243–263.
Johnson, Timothy R. 2009. "The Digital Archives of Justices Blackmun and Powell Oral Argument Notes." Available at http://users.polisci.umn.edu/~trj/oanotes93.php.
Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus, & Giroux.
Kahneman, Daniel, and Amos Tversky. 1982. "On the Study of Statistical Intuitions." *Cognition* 11(2): 123–141.
Kastellec, Jonathan P. 2020. "Race, Context, and Judging on the Courts of Appeals: Race-Based Panel Effects in Death Penalty Cases." *Justice System Journal*: 1–22.
Kropp, Sabine. 2010. "German Parliamentary Party Groups in Europeanised Policymaking: Awakening from the Sleep? Institutions and Heuristics as MPs' Resources." *German Politics* 19(2): 123–147.

https://doi.org/10.1017/jlc.2024.2 Published online by Cambridge University Press

Lenz, Gabriel S., and Chappell Lawson. 2011. "Looking the Part: Television Leads Less Informed Citizens to Vote Based on Candidates' Appearance." *American Journal of Political Science* 55(3): 574–589.

MacCallum, Robert C., Shaobo Zhang, Kristopher J. Preacher, and Derek D. Rucker. 2002. "On the Practice of Dichotomization of Quantitative Variables." *Psychological Methods* 7(1): 19–40.

Maroney, Terry A. 2015. "Why Choose? A Response to Rachlinski, Wistrich and Guthrie's 'Heart Versus Head: Do Judges Follow the Law or Follow Their Feelings?.' " *Texas Law Review* 93(1): 317–330.

Martineau, Robert J., Kent Sinclair, Michael E. Solomine, and Randy J. Holland. 2005. *Appellate Practice and Procedure*. St. Paul, MN: West Academic Publishing.

Meissner, Christian A., and John C. Brigham. 2001. "Thirty Years of Investigating the Own-Race Bias in Memory for Faces: A Meta-Analytic Review." *Psychology, Public Policy, and Law* 7(1): 3.

Mintz, Alex. 2004. "How Do Leaders Make Decisions? A Poliheuristic Perspective." *Journal of Conflict Resolution* 48(1): 3–13.

Monin, Benoît. 2003. "The Warm Glow Heuristic: When Liking Leads to Familiarity." *Journal of Personality and Social Psychology* 85(6): 1035–1048.

Nelson, Michael J., and Lee Epstein. 2021. "Human Capital in Court: The Role of Attorney Experience in U.S. Supreme Court Litigation." *Journal of Law and Courts*. doi: 10.1086/714577.

Nordholm, Lena A. 1980. "Beautiful Patients Are Good Patients: Evidence for the Physical Attractiveness Stereotype in First Impressions of Patients." *Social Science & Medicine. Part A: Medical Psychology & Medical Sociology* 14(1): 81–83.

O'Brien, David M. 1981. *Judges on Judging: Views from the Bench*. Washington, DC: CQ Press.

Olivola, Christopher Y., Dustin Tingley, and Alexander Todorov. 2018. "Republican Voters Prefer Candidates Who Have Conservative-Looking Faces: New Evidence from Exit Polls." *Political Psychology* 39(5): 1157–1171.

Posner, Richard A. 2010. *How Judges Think*. Cambridge, MA: Harvard University Press.

Quinn, Joseph W. 2012. *A Judge's View: Things Lawyers Do That Annoy Judges; Things They Do That Impress Judges*. Toronto, ON: Family Law Institute.

Rachlinski, Jeffrey J., Andrew J. Wistrich, and Chris Guthrie. 2015. "Can Judges Make Reliable Numeric Judgments: Distorted Damages and Skewed Sentences." *Indiana Law Journal* 90(2): 695–739.

Rachlinski, Jeffrey J., Chris Guthrie, and Andrew J. Wistrich. 2007. "Heuristics and Biases in Bankruptcy Judges." *Journal of Institutional and Theoretical Economics* 163(1): 167–186.

Rachlinski, Jeffrey J., Sheri Lynn Johnson, Andrew J. Wistrich, and Chris Guthrie. 2009. "Does Unconscious Racial Bias Affect Trial Judges?" *Notre Dame Law Review* 84(3): 1195–1247.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1): 41–55.

Scalia, Anontin, and Brian A. Garner. 2008. Making Your Case: The Art of Persuading Judges. *St*. Paul, MN: Thomson West.

Segal, Jeffrey A., and Harold J. Spaeth. 2002. *The Supreme Court and the Attitudinal Model Revisited*. Cambridge: Cambridge University Press.

Spamann, Holger. 2015. "Empirical Comparative Law." *Annual Review of Law and Social Science* 11: 131–153.

Spamann, Holger, and Lars Klöhn. 2016. "Justice is Less Blind, and Less Legalistic than We Thought: Evidence from an Experiment with Real Judges." *Journal of Legal Studies* 45(2): 255–280.

Stern, Robert L., Eugene Gressman, Stephen M. Shapiro, and Kenneth S. Geller. 2002. *Supreme Court Practice*. 8th ed. Arlington, VA: Bureau of National Affairs, Incorporated.

Stewart, John E. 1980. "Defendant's Attractiveness as a Factor in the Outcome of Criminal Trials: An Observational Study 1." *Journal of Applied Social Psychology* 10(4): 348–361.

Stockemer, Daniel, and Rodrigo Praino. 2017. "Physical Attractiveness, Voter Heuristics and Electoral Systems: The Role of Candidate Attractiveness under Different Institutional Designs." *British Journal of Politics and International Relations* 19(2): 336–352.

Thorndike, Edward L. 1920. "A Constant Error in Psychological Ratings." *Journal of Applied Psychology* 4(1): 25–29.

Todorov, Alexander, Anesu N. Mandisodza, Amir Goren, and Crystal C. Hall. 2005. "Inferences of Competence from Faces Predict Election Outcomes." *Science* 308(5728): 1623–1626.

Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185(4157): 1124–1131.

Wasby, Stephen L. 1981. "Oral Argument in the Ninth Circuit: The View from Bench and Bar." *Golden Gate University Law Review* 11(1): 21–79.

Wedeking, Justin. 2010. "Supreme Court Litigants and Strategic Framing." *American Journal of Political Science* 54(3): 617–631.

Weeden, Jason, and John Sabini. 2007. "Subjective and Objective Measures of Attractiveness and Their Relation to Sexual Behavior and Sexual Attitudes in University Students." *Archives of Sexual Behavior* 36 (1): 79–88.

Weinshall, Keren, and Lee Epstein. 2020. "Developing High-Quality Data Infrastructure for Legal Analytics: Introducing the Israeli Supreme Court Database." *Journal of Empirical Legal Studies* 17(2): 416–434.

Zebrowitz, Leslie A., and Susan M. McDonald. 1991. "The Impact of Litigants' Baby-Facedness and Attractiveness on Adjudications in Small Claims Courts." *Law and Human Behavior* 15(6): 603–623.

Cite this article: Waterbury, Nicholas W. 2024. "Justice Isn't Blind: Attorney Attractiveness and Success in US Federal Court." *Journal of Law and Courts*, 1–22, doi:10.1017/jlc.2024.2